

The US budget crisis can only recur

It is not an accident that the US Congress has been battling with President Bill Clinton over this year's budget. Both are under global pressure. Each must take care that their compromise does not wreck science.

It is easy to make fun of the United States because the administration and the Congress have yet again failed to reach a compromise on the federal budget for the financial year that began on 1 October. What a way to run a great country is the common jibe. It is not, after all, that these two organs of the US government have been short of time; budgets of one origin (the White House) or another (the Capitol) have been on the table almost since the beginning of the year.

All that is beside the point. The Congress and the administration disagree about the scale on which public funds can be spent, this year more sharply than for decades. The Congress, or its Republican majority, is simply exercising its unique right under the US Constitution to determine what the administration can legitimately spend. The trouble is worse this year because the President needs not just a renewal of the continuing resolution (allowing the government to spend at last year's rate), but an increase of the ceiling on the public debt. Yet even Americans inconvenienced by the past week's crisis will proclaim it as a proof that democracy is alive and well in the United States. And they are of course correct.

Nevertheless, the United States would find the inevitable compromises that will have to be made more palatable if they saw them in a wider context. The big crunch, yet to come, is the great Welfare Bill on which the Republicans in both houses of the Congress are still working. Different versions of the bill have already been passed; now they have to be reconciled. Unfamiliarly, the ruling party knows that there is nothing to stand between whatever compromise it reaches and enactment of the bill except a presidential veto. It suits politicians on both sides to represent that struggle, not to mention the disagreement over a balanced budget (five, seven or more years ahead) to which end the welfare bill is part of the means, as a dispute between the good guys (themselves) and the bad guys (their opponents), but that is not the whole story. Even the United States, for long the largest of all island states, has been overtaken by globalization.

Mobility

This is how it works. Every democratic government's ambition is to keep its voters in the state of prosperity they believe to be their due, which evidently restricts the scope for increasing taxes. Even worse (as Sweden found

to its cost a decade ago), people can often now move to where taxes are lower, thereby creating a kind of international standard in top tax rates. While international mobility persists, 'soaking the rich' is impracticable. But nowhere is everybody rich. Everywhere there are some people without jobs, many of whom, in modern and compassionate societies, need welfare. How is that cost to be met?

Competition

Governments' freedom is again curtailed by international considerations. One solution is to print extra money to pay welfare and other social running costs, but the global financial markets will no longer tolerate that; instead, they quickly switch short-term investments from currencies heading for inflation, and the governments responsible have to pay above the odds to rub along from day to day. So why not get the people without jobs back to work, so that instead of consuming public funds they will enlarge them through the taxes they pay at work? Sadly, even that constructive course of action is now hemmed in by international competition. Increasingly, it is pointless to create jobs in industries which are not internationally competitive. The international markets are not short of capital, which will find its way to the places where things of particular kinds can be made most cheaply.

The implications are (or should be) sobering, and not just for the United States. Industrialized governments are pinioned from three directions: they cannot raise top tax rates (very much), they cannot run persistent budget deficits except in special circumstances (as when, in the Reagan years, Japan reinvested its trade surplus in the United States) and they can sustain their people in a specified state of well-being only by making sure that there are innovative occupations in which they can be engaged, perhaps only for as long as it takes other people to catch up. Add to that the extra worry that most rich countries have ageing populations whose members have been promised social support and even pensions by their governments, and it can be seen that this week's row in Washington is unavoidable.

It is no accident that governments everywhere are doing what they can to cut their spending bills. That is the post-election message of President Jacques Chirac of France to those who elected him on a different ticket in the summer; he may say and even believe that his



purpose is to comply with the requirements of the European Monetary Union planned for some time soon, but globalization would have left him no alternative. Britain's Chancellor of the Exchequer will be doing much the same later this month (although he will have to say that his purpose is to reduce taxes, for his party does not like to think that it is a pawn in the hands of foreigners).

The danger in this new but inescapable fashion for reduced government spending is that, without an open understanding of why it is necessary, it can do more harm than good. Governments like that of the United States have been brooding about 'competitiveness' for a quarter of a century, wringing their hands over the rundown of once thriving industries such as shipbuilding and shoe manufacture. They have migrated elsewhere, and will not come back. But the United States remains as competitive as anybody could ask in fields such as those represented by Intel, Microsoft and AT&T. The immediate problem is to create the conditions in which employment in those fields will grow substantially, which means high-level training. Further ahead is the need to capture the next wave of innovation, which means research. But both of those are causes covered by discretionary spending, which can and probably will be cut. And what is true for the United States is true almost everywhere else. As might be expected, globalization works everywhere. □

Europe's black sheep

EU member states spending too little on research and development need carrots as well as sticks.

ANTONIO Ruberti, the member of the European Commission responsible for research until the end of last year, has a good idea (see page 226): let European governments wishing to join international European collaborations first put their own houses in order by demonstrating that they are already investing enough in research to benefit from the collaboration they wish to join. The idea is a good one because there is an urgent need to bring about a more uniform pattern of spending on research in the member states of the European Union (EU). As things are, Greece (where Aristotle used to live) is at the bottom of all lists; total spending on research and development runs at about 0.5 per cent of gross domestic product. Portugal (0.7 per cent) is hardly better, Ireland (1.1 per cent) is next from the bottom.

Membership in itself evidently does not create the climate in which all governments pull their weight in research. Of course, comparisons between the performances of the three black sheep with those of the larger member states are to some degree unfair. Both Britain and France, for example, spend substantial sums on military research, while Germany and France spend substantially on space research in various ways — activities that would not make sense in the smaller European states. But that is only a partial defence of the neglect of science

by the EU's smaller members. For one thing, it denies the broader community the potential benefits of their talent, while in the long run it will impoverish their own people by robbing them of modern skills.

Ruberti's condition for collaboration is therefore something like a stick with which the three black sheep might be beaten. Although it is unlikely that Ireland will ever be an enthusiastic member of the European Space Agency, or of CERN, there is ample reason to expect that young men and women from all three countries would (given a start) find their way to Heidelberg (and the European Molecular Biology Laboratory, whose current director is Greek) or Munich (the headquarters of the European Southern Observatory). It makes sense to ask that the black sheep governments should provide the means by which the people concerned can be trained in the first place and then can be enabled to spend their time productively at work when their collaboration is complete. Ruberti's device is therefore a little like a stick with which to beat the laggard governments, even if into pursuits that serve their own long-term interests.

But if there is to be a stick of some kind, there should also be a few carrots. What might those be? The example of Ireland suggests stratagems that could help. It is too easily forgotten that Ireland is where Schrödinger settled after the Second World War, when the Dublin Institute for Advanced Study was active in a variety of modern fields, cosmic rays for example. But all trace of that endeavour had vanished by the mid-1960s, submerged by Celtic studies of various kinds. Trinity College Dublin, now a constituent of the University of Ireland (where Fitzgerald of the Lorentz-Fitzgerald contraction worked), has kept at research on a broader front, but it is plainly difficult to build up large research groups when the intellectual hinterland is so thinly populated. Yet from Cork (with excellent chemistry) to Armagh (with its observatory), there is no reason to suppose that Ireland (a land of scholars) is indifferent to research. The need is for people.

That is the carrot the EU could help to provide. As things are, the Human Mobility Programme (whose rules have now been sensibly rewritten — see *Nature* **378**, 118; 1995) will no doubt help many young people from the smaller countries to spend time at established laboratories elsewhere, only to find themselves at a loose end when their fellowships are up. Ideally the movement should be in the opposite direction. But the European Commission was given a formal competence in education by the Maastricht Treaty. While all member-states hold firmly to the view that their own educational systems are near-perfect, and also jealously guard their right to subsidiarity, the commission is unlikely to be allowed to do much in the field. The creation of a few handfuls of European faculty posts at carefully chosen universities, coupled to research support, could lead to the emergence of research groups of distinction. And that could transform the climate, so that the black sheep were no longer black. Is that not worth a try? □

Howard Hughes to shift focus of support away from institutions

Washington. The Howard Hughes Medical Institute (HHMI) is to stop funding research at US institutions where HHMI-supported investigators lose their grants or move to different institutions, according to officials at the Maryland-based research organization.

The change in policy will allow the organization to withdraw from the partnerships it has formed with universities when it stops funding particular investigators at these institutions. HHMI views this as a logical step in a strategy that it has been following since 1986 of shifting away from making awards only to investigators at universities with which it is already in partnership, and making them to the best-qualified applicants at any institution.

HHMI, which uses the fortune bequeathed by the late eccentric aviation billionaire Howard Hughes to fund research in molecular biology, is the largest non-government supporter of university science in the United States. Last year it spent \$280 million on its investigators, and a further \$50 million on various grants.

Purnell Choppin, the president of HHMI,

said last week that the money that up to now has supported such investigators will revert to a central pool, from where it can be redirected to new investigators anywhere in the United States who are successful in open competition for HHMI awards.

The move will make more HHMI money available for fresh investigators, including those at institutions where the institute has no presence. But the very low level of attrition means it will not lead to a deluge of new grants. Bob Potter, head of communications at HHMI, says the next round of new awards will probably be made in late 1996 or early 1997, adding "ten or twelve" investigators to the current pool of 280. HHMI spends an average of \$1 million a year supporting each member of this elite group.

Part of the rationale for HHMI's previous

policy was a wish to prevent all of its money from following the best researchers into a tiny handful of molecular biology departments. But Choppin says that its recent, open competitions have not caused this to happen; in 1994, when HHMI made 44 new awards in open competition, it added ten new institutions, and its investigators are now spread across 62 institutions.

Choppin says that HHMI researchers may also get more freedom to change institutions and take HHMI money with them. "At the moment we consider this on a case-by-case basis, and in most cases, they are not free to move," he says. "But we may reconsider that."

Colin Macilwain

Columbia takes over reins at Biosphere 2

Washington. Columbia University's Lamont-Doherty Earth Observatory at Palisades, New York, is to take over the entire operation of Biosphere 2, the Arizona ecology laboratory built by Texan millionaire Ed Bass, from the beginning of next year, according to officials from the university.

Columbia's take-over of Biosphere 2 completes a period of transition during which Bass ousted the ideologically committed co-founders of the laboratory, and appointed an outside committee of advisers, chaired by Michael McElroy of Harvard University's Earth sciences department, to restore its battered scientific reputation (see *Nature* 368, 577 & 370, 495; 1994).

At the end of next month, Biosphere 2's acting chief executive, Stephen Bannon, will be replaced temporarily by Taro Takahasi, an associate director of Lamont-Doherty. A search will then be made for a permanent director, who could be in place by April.

Michael Crow, vice provost of Columbia University, says that up to eight additional groups of scientists would start working at Biosphere 2, and that a small, self-financing college would open there. "There will be a greatly expanded scientific presence, with people from Lamont-Doherty and from other institutions," he says.

Bass has spent an estimated \$150 million building the huge, hermetically-sealed Biosphere 2 greenhouse in the Arizona desert. The project drew much derision in its early days, when staff tried to survive inside the greenhouse for several months without external support. But it is being increasingly accepted by ecologists and others as a potentially powerful tool for investigating the complex relationship between the earth, the atmosphere and living things.

C. M.

Dirac given his place next to Newton

London. A mathematical equation that helped to seal the fate of Newtonian mechanics was unveiled next to the tomb of Sir Isaac Newton at a special service at Westminster Abbey in London at the beginning of this week.

The equation is carved on a 60-centimetre-square slate plaque (right) in memory of the theoretical physicist Paul Dirac. Dirac was Lucasian Professor of Mathematics at the University of Cambridge for 37 years, a chair once occupied by Newton. During the service, an address was given by the current occupant of the chair, Stephen Hawking.

Dirac, born in Bristol and one of the founders of quantum mechanics, developed the mathematics by which Einstein's theory of relativity was brought into the quantum mechanical theory of electrons. The so-called Dirac Equation, $i\gamma_\mu \partial_\mu \psi = m\psi$, encapsulates the concept that an electron can have two different energy states, one positive and one negative.

Because electrons are negatively charged, Dirac's logic suggested that a similar particle with a positive charge ought to exist.

Protons were ruled out, as they are 1,836 times as massive as electrons. In 1930, Dirac suggested that electrons must have a positive twin. This

idea of antimatter was confirmed in 1932. The following year Dirac shared the Nobel prize for physics with Erwin Schrödinger.

The plaque, made in the Cambridge workshop of the stonemason David Kindersley, represents the first inscription of a mathematical equation inside Westminster Abbey. But the unveiling, performed by Sir Michael Atiyah, president of the Royal Society and Master of Trinity College, Cambridge, had a further twist: whereas Newton, whose mechanics painted a deterministic picture of the world, was a Christian, Dirac, whose sub-atomic world was more difficult to predict, was an obdurate atheist.

Ehsan Masood

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Michael Mann Photograph

France backs off promise to boost UN AIDS programme

Paris. The French government has reneged on a promise to donate FF100 million (US\$20 million) to the new joint United Nations Programme on HIV/AIDS — known as UNAIDS — as part of broader cuts in foreign aid. The decision has provoked widespread protest in both France and elsewhere, as well as calls for it to be reversed.

The funding promise was made at last year's 'AIDS Summit' in Paris by then prime minister Edouard Balladur. Its apparent revocation has reinforced the views of sceptics that the meeting was largely a political stunt, as well as concern that support promised by political leaders at the meeting for the fight against AIDS would become a "long list of good intentions" that may never be translated into action (see *Nature* 372, 308 & 493; 1994).

Peter Piot, for example, the director of the UNAIDS programme, created late last year by five UN agencies and the World Bank as a replacement for the Global AIDS Programme of the World Health Organization (WHO), has described the French decision as "very worrying and profoundly disappointing".

One immediate consequence, according to Piot, will be the cancellation of series of planned programmes in developing countries. The full impact will not be known until later this week, when the programme's coordinating board meets to discuss its planned budget of \$120–140 million over the next two years.

One WHO official says that UNAIDS is confident that the programme will obtain the necessary commitments from donor countries. But the French decision — which WHO still hopes the French government will reverse in time for this week's meeting — will inevitably add to concern that UNAIDS may not obtain the financial and political support needed to accomplish its ambitious goals, particularly as it comes at a time when many industrialized countries are reducing their foreign aid budgets.

The pressure group AIDES, as well as several other French patient groups, have reminded the government that Balladur himself declared at the Paris summit that "85 per cent of people infected with the AIDS virus live in disadvantaged countries", and says that "France would not be France any more, if it came to accept this situation".

Such protests have also been taken up forcefully by the newspaper *Le Monde*, which in a reference to the AIDS summit, declared in a leading article that "the deception is in line with the expectations created".

D. B.

Geneticist quits in protest at 'genes and violence' claim

Paris. The French president-elect of the Behaviour Genetics Association has resigned in protest at the refusal of the association's executive committee to expel a past-president for having given a speech in which he claimed that there was a racial basis for different murder rates between US cities.

The speech was given at the 25th annual meeting last May of the 450-member association by Glayde Whitney, a psychologist at Florida State University. In the speech, Whitney argued that genetic inheritance may play a major role in determining behaviour and condemned what he described as the "left-liberal Marxist dogma of environmental determinism".

Presenting data which, he claimed, demonstrated a strong correlation between murder rates in US cities and the proportion of blacks, Whitney argued that "some, perhaps much" of this could be the result of genetic differences related to intelligence, "lack of empathy", aggression and "impulsive lack of foresight". He added: "I know of no environmental variable that accounts for more of the variation [than the proportion of blacks in the population]".

Many members of the both the association and the executive committee have since denounced the speech as "offensive". For example, the current president, James Wilson, of the University of Colorado's Institute for Behavioural Genetics, has described the address as hurtful both to black members of the audience and to black staff at the function, while the association's journal, *Behaviour Genetics*, has refused to publish the address.

But Whitney has now published his address in the journal *Mankind Quarterly*, accompanied by a preface entitled "Ideology and Censorship in Behaviour Genetics" in which he says that his speech was intended to show that "ideologically-based dogma and taboo" are hampering research in behavioural genetics.

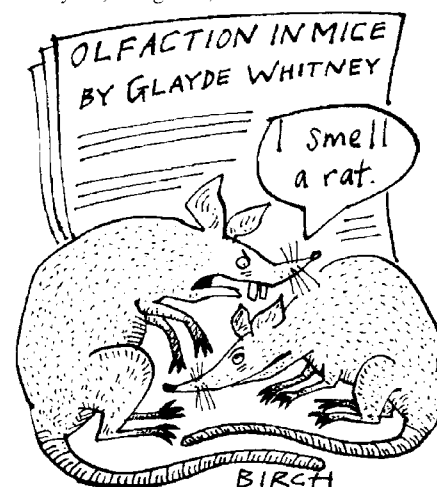
Whitney, who is renowned for his work on olfaction in mice, was not available for comment last week. But in his preface, he described the reaction to his speech as a "parody of political correctness", and criticized members of the association's executive committee for acting in the "finest Lysenkoist tradition".

Despite such criticism, the executive committee has decided to take no action against Whitney. According to the association's secretary, Andrew Heath, of the department of psychiatry at the Washington School of Medicine in St Louis, Missouri, any action to penalize Whitney would be contrary to freedom of speech.

Furthermore, the committee, after consultations with former board members, cancelled a meeting scheduled to take place last week at which the issue of whether it would be "appropriate" to expel Whitney from the executive committee was to have been discussed. Holding the meeting would also only have "inflamed the situation", says Wilson, who himself had earlier urged Whitney to resign.

According to Heath, following these consultations, the committee concluded that Whitney had showed "poor judgement". But a majority of its members also felt that it could not "take the risk" of compromising the right of scientists to free speech. "The price we pay for this freedom is that sometimes we have to listen to things that might offend us as individuals," says Heath.

But Pierre Roubertoux, who was to have taken over as the association's president next year, disagreed, and announced that he



was no longer prepared to take up the new post — Roubertoux heads the Centre National de la Recherche Scientifique (CNRS)/University of Paris V René Descartes Laboratory of Genetics, Neurogenetics and Behaviour.

In his resignation letter, Roubertoux argues that freedom of speech should apply to a talk given as one individual's point of view at a symposium, where it could be discussed and if necessary challenged. But he claimed that, as the address was given by an elected officer, it appeared to be "the official view of the association".

Roubertoux, who walked out during Whitney's speech, says he has since apologised on the association's behalf to black waiters working at the function. "I refuse to be on the same board as Whitney," he adds. Wilson describes Roubertoux's resignation as a "heavy loss to the board."

Declan Butler

Russian fund gets off the ground — at last

Washington & Moscow. The US Civilian Research and Development Foundation (CRDF), set up last August to fund joint projects between scientists in the United States and the countries of the former Soviet Union (FSU), last week put out its first call for proposals.

The CRDF now has \$10 million in hand, thanks to the efforts of the US government and the philanthropist George Soros, who settled a dispute about support for the non-profit fund earlier this year (see *Nature* 375, 170; 1995). The foundation also has promises of \$5 million more from Russia, and a possible \$1.5 million from the Ukraine.

Only \$6 million of those funds will be used for this first round of competitive grants. Successful research teams will receive between \$10,000 and \$80,000 for a two-year grant period. Any area of civilian research and development will be considered for support, but projects with a commercial or entrepreneurial emphasis will be favoured, as will FSU researchers who formerly worked on defence projects.

All applications will be reviewed by CRDF panels of scientists and engineers, and the winners will be selected by next July.

Several factors distinguish the CRDF grants from previous aid programmes to Russian scientists, according to Peter Raven of the Missouri Botanical Garden, the chairman of the foundation's board of directors. First, although the grants are relatively modest, they represent the first US government money awarded for research in the FSU.

Second, the grants cover both basic and applied research, without any distinction being made between the two. And third, every team must include both an FSU scientist and a US scientist, a condition intended to build partnerships between researchers and foster intellectual communication.

Few pretend that the CRDF money will, on its own, be sufficient to save Russian science. Irving Lerch of the American Physical Society, which has taken an active role in fostering collaboration with FSU researchers, says that there are too many scientists looking for work, too much politi-

cal infighting in Russia and too many economic and cultural hurdles for the problem to be solved overnight.

But Lerch says that the CRDF "can have a salutary impact if it's done cleverly, and if it makes some effort to coordinate with the other programmes going on worldwide", such as the European Union's INTAS (see *Nature* 375, 345; 1995) and the US Industry Coalition (USIC), funded by the Department of Energy, which will provide \$35 million to support collaboration with researchers in DoE laboratories.

Raven also admits that "the [CRDF] money is not going to go very far", given that there are 600,000 former-Soviet scientists and engineers in need of support. But he hopes the strategy of targeting commercially useful research will benefit FSU economies, and that these will in turn improve the situation for researchers in those countries.

The foundation's board will soon consider whether to create a separate, \$2-million grant programme limited to cooperation in industrial research and development. But Raven says that this is still unexplored territory, and that, whatever the board decides, the CRDF can be only a small player in the broad game of building East-West industrial partnerships.

The CRDF was first proposed in US legislation sponsored by Congressman George Brown of California, then-Senator Al Gore of Tennessee, and Senator Joseph Lieberman of Connecticut. The resulting authorization to create the CRDF was passed by the US Congress and signed into law in 1992.

The CRDF's initial funding comes from a \$5-million allocation from the Department of Defense's Nunn-Lugar programme, designed to promote the demilitarization of the FSU, and a matching \$5-million gift to the US National Science Foundation from Soros.

Tony Reichhardt & Carl Levitin

Political muscle for French science

Paris. Last week's surprise reshuffle of the French government brought the removal of both Elisabeth Dufourcq (RPR), the secretary of state for research, and Elisabeth Hubert (RPR), the health minister, as well as the demotion of François Fillon (RPR), previously minister for information technology, who becomes a junior minister responsible for the Post Office, Telecommunications and Space.

Dufourcq is replaced by François d'Aubert (UDF-PR), who was secretary of state for the budget in the previous government. The move is widely seen in Paris as stemming less from unhappiness with Dufourcq — who despite her lack of political experience, was considered otherwise to be doing a reasonable job — and more from the need to find a new position for d'Aubert, who is said to have been in conflict in his previous post with the finance minister, Jean Arthuis (UDF-CDS).

It is too soon to predict d'Aubert's likely impact on research. But observers point out that he is a political heavyweight, being vice-president of the UDF group — one of the main groups that make up the conservative coalition headed by Alain Juppé, the prime minister — in the

National Assembly. D'Aubert's experience in the finance ministry could make him either an effective defender of research budgets, or alternatively an aggressive proponent of deep budgetary reforms of the research system.

Dufourcq and Hubert were two of eight women removed from the government in a reshuffle that reduced the total number of ministers and junior ministers from 41 to 32. The replacement of one of the shortest-lived governments in recent history is aimed at reversing the plummeting popularity of Juppé and Jacques Chirac, the French president, by replacing the previous floundering government with a tighter and more experienced team.

The health ministry is absorbed within a super-ministry of employment and social affairs, under Jacques Barrot (UDF-CDS), whose main task will be to reduce the social security deficit. Hubert, a close ally of Chirac, was sacked for indiscretion and unreliability. Her health portfolio goes to a secretary of state, Hervé Gaymard (RPR).

The demotion of Fillon, who led the French delegation to last month's space ministers meeting of the European Space Agency (see *Nature* 377, 667–668; 1995), and was the cabinet minister responsible for research and universities in the previous government, follows his poor handling of plans to privatize France Telecom.

But the armed forces are said to have supported continuing the responsibilities of Fillon, who is a defence expert, for the space portfolio.

Declan Butler

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

D'Aubert: which way
will he use his skills?

Gamma

Montreal wins bid for biodiversity centre

London. Montreal in Canada has been chosen as the location of the secretariat of the UN Convention on Biological Diversity, signed at the Earth Summit in Rio de Janeiro in 1992. The decision was announced on Monday at the beginning of the second week of the annual conference of parties to the convention, which is being held in the Indonesian capital Jakarta.

Montreal was chosen in preference to Nairobi — which houses the headquarters of the United Nations Environment Programme — and Geneva, which is expected to continue to house the secretariat temporarily for a further year. The fourth candidate was Madrid, the Spanish capital.

Collaboration rules should reflect investment

Florence. Any new country wishing to join a major European collaborative research venture should be required to demonstrate that it is making sufficient domestic investment in science to be able to benefit directly from such collaboration, Antonio Ruberti, the former European commissioner for research, suggested last week.

At the same time, new rules need to be developed for scientific collaboration with major industrial nations — in particular the United States and Japan — to allow Europe to benefit from the strengths it has built up over the past several decades, and ensure that it is no longer treated as a 'junior partner' in any global co-operative venture.

Both suggestions emerged at the end of a three-day meeting held in Florence to discuss the lessons to be drawn from almost 50 years of experience in European scientific collaboration — starting with those efforts, such as the building of the European Laboratory for Particle Physics (CERN), which have their roots in the reconstruction of the continent at the end of the Second World War — and to suggest how such lessons might guide future plans.

Furthermore, both reflect a growing shift in the attitude of European politicians towards collaborative ventures. There is now a reduced emphasis on the potential political value of such ventures, and increased concern for their cost-effectiveness, judged primarily in scientific and technological terms.

The meeting provided some clear divergences of view on both the lessons of the past, and their implications for the future. Organized by the European University Institute, and funded by a grant from the European Commission, it provided an opportunity for several historians of science to provide evidence of the extent to which, during this period, collaboration has been driven as much by political as scientific considerations.

In some instances, the historians' analysis received wide support. There was general endorsement, for example, for a statement by John Krige, one of the organizers of the conference who is currently writing a history of the European Space Agency (ESA), that post-war scientific collaboration in Europe must take the desire to both compete and collaborate with the United States as "a fundamental point of reference".

But several other conclusions from the historians received a more critical response from the second main group represented at the Florence meeting, namely scientists and government officials responsible for specific collaborative efforts.

There were, for example, some sceptical responses from those reluctant to emphasize the links between science and politics to the suggestion of Krige and his colleague

Dominique Pestre, a historian of post-war French physics, that the construction of CERN had been implicitly endorsed by Europe's military establishment because of its potential, if indirect, contribution to new nuclear technologies.

Yet whatever the disagreements between the historians and practitioners of collaboration, consensus emerged on several key points. One was that there is "no optimal model" for scientific collaboration, the



Hubert Curien (centre), the former French research minister, addresses last week's meeting, flanked by Bertel Haarder (left) of Denmark and Antonio Ruberti (right) of Italy.

nature of successful institutional arrangements depending highly on the specific goal being pursued.

A second point of virtually unanimous agreement was that a high price for European-wide collaboration is paid by scientists in terms of the time-consuming bureaucratic hurdles over which funding applications for joint research projects have to pass.

Some blamed this excessive bureaucracy on the enthusiasm of politicians for policies of *juste retour* — the idea that an individual country's contribution to a particular project should be reflected in the share of industrial contracts it receives — and the resulting need felt by administrators to be able to justify all decisions in this context.

Others, such as Jean Pierre Contzen, director general of the commission's Joint Research Centre in Ispra, Italy, highlighted the way restrictions imposed for political reasons can reduce the flexibility of European Union (EU) institutions to react flexibly to new challenges.

Indeed, there was widespread support for a call from Sir John Kendrew, former director general of the European Molecular Biology Laboratory (EMBL), for a serious study of the relationship between science and bureaucracy. "Many scientists think that there has become too much of it," he said.

A third point of apparent consensus was a feeling that the rules for membership of European-wide projects, many of which were originally established to deal with the relatively small number of countries making up western Europe, need to be reassessed in the light of growing demands for member-

ship from the newly independent states of eastern Europe.

Tensions on this issue have been increasing steadily over the past few years, originating not in opposition to the principle that such states be permitted to join established collaborative projects, but in the distortions that their presence may cause to established procedures, ranging from voting rules based on the concept of 'one country, one vote', to the application of *juste retour* principles.

Such pressures are being felt intensely, for example, by CERN. After the collapse of communism, CERN agreed to admit a number of new eastern European countries on relatively generous terms. But CERN's director general, Christopher Llewellyn Smith, told the meeting that the pressures of expanded membership, coinciding with increased demands from member states for a distribution of service contracts that reflects their financial contribution to the agency, meant it was unlikely to be equally generous to those now knocking at the door.

Some suggested that membership rules adopted by bodies created more recently, and thus able to base their practice on the experiences of their predecessors, offered possible solutions. Yves Petroff, for example, director general of the European Synchrotron Radiation Facility in Geneva, pointed to the success of a rule that smaller countries could only become members as part of a consortium contributing a total of at least four per cent of the overall costs.

Ruberti's proposal was even more radical, namely that the government of any country seeking to participate in joint European projects be required to demonstrate a domestic commitment to increased funding for science before being allowed to do so. Ruberti's proposal reflects the principle of 'convergence' already adopted as the basis of future monetary union in the EU, namely that this be restricted to those able to meet a fixed set of economic criteria.

"We have to take on board the different levels of research efforts in different countries, and to become aware of the need for convergence in national efforts," said Ruberti. Without this, he added, it was difficult to endorse cooperation between countries "which sometimes differ by a factor of six in terms of the proportion of national product spent on research."

Several other participants at the meeting supported the idea, pointing out that it could be used to encourage poorer countries within the EU to use the so-called 'structural funds', intended by Brussels to help build up their economic infrastructure, to strengthen their scientific and technological capabilities, rather than using such funds solely for more conventional infrastructure projects, such as road-building. How Brussels reacts remains to be seen.

David Dickson

Japanese science law could bring control as well as cash

Tokyo. A new law was passed by the Diet, Japan's parliament, last week, designed to ensure adequate funding of public sector research by making such funding a legal responsibility of the government. The law requires a five-year basic plan for science and technology to be drawn up, and the government will have to report its progress annually to the Diet. It also urges the government to promote increased contact with researchers in other countries.

The Basic Science and Technology Law had been proposed by Omi Koji of the Liberal Democratic Party (LDP), the coalition government's largest party. Koji is a former official at the Ministry of International Trade and Industry (MITI) and one of a group of LDP members who have been pushing for increased support for science and technology (see *Nature* 378, 8; 1995).

The law has been welcomed by officials at MITI and the Science and Technology Agency (STA). But some scientists fear that it could lead to more direct control of academic research by the government. Similar concern, expressed at the time by the Science and Technology Council — Japan's leading science advisory body, made up of representatives from government, academic institutions and industry — helped to defeat a similar bill in 1968, although it was also opposed by the Ministry of Finance.

The new law does not specify the precise level at which the government should fund science. But Koji expects that it will lead to increased funding for science in the 1996 budget across all government agencies and

ministries. He claims that the Japanese government, unlike some of its competitors, such as the United States and Germany, contributes a relatively small proportion of national spending on research and development, and that this should be increased for the sake of the economy.

Atsushi Kawashima, the director general for general affairs at MITI's Agency of Industry, Science and Technology (AIST), says the law is expected to raise the awareness of Japanese legislators of the importance of science and technology. He believes that funding increases will be specified in the basic science plan which the government is now required to draft.

Reaction to the law in the academic community has reflected a combination of support and wariness. Itaru Watanabe, for example, professor emeritus at Keio University, says that while the law is likely to benefit industry-oriented research, this may be at the expense of basic research. Junjiro Kanamuri, president of Osaka University, welcomes the new legislation, but emphasizes that academics need to be consulted in drafting of the basic science plan.

Akita Arima, former president of Tokyo University and current president of the Institute of Physical and Chemical Research (RIKEN), is pleased with the new law's implied promise of increased funding, as well as the greater 'internationalization' of Japanese science. But, like Watanabe, he hopes that funding will not be restricted to areas with commercial applications.

Stephen Barker

European Commission seeks clearer goals for space research

Munich. Edith Cresson, the European Union (EU) commissioner for science, last week claimed that Europe lacks a long-term political vision in space research, despite the economic and scientific importance of space activities.

Speaking at the European Space Forum, a meeting of policy-makers and space technologists organized jointly by the European Commission and the European Parliament, Cresson criticized the lack of coordination of programmes that make use of such activities within the commission.

Space had appeared to be so unimportant to the commission "that it hadn't dreamt of having a director in charge of the section or having a budget for coordination [of space utilization programmes]," she said.

Cresson claimed that national space agencies, as well as the European Space Agency (ESA), have promoted the development of cutting-edge technologies, but that European industry continues to lag behind that of the United States and Japan, particularly in the use of satellite data.

The commission had built up its own programmes in the use of satellite technologies, particularly in telecommunications, navigation and Earth observation. But Cresson argued that these needed to be better coordinated in order to create new markets.

In response to pressure from the European Parliament, concerned at its own lack of direct control of European space policy as this has historically been in the hands of ESA, an intergovernmental body, Cresson is now preparing a report to be presented both to the parliament and to the council of ministers next spring.

The report is being drawn up jointly with Martin Bangemann, the commissioner for industry and telecommunications, and will analyse the commission's programmes involving the use of satellite data and suggest how they might be developed.

Alain Pompidou, a French member of the European Parliament, has already suggested that the commission should set up a task force to develop and promote an EU space policy. This would not be intended to replace ESA, he says, "but it would focus its efforts on specific policy goals, such as meeting users' demands".

Cresson's announcement that the report on promoting the use of space-based data will be ready by the spring is seen as over-optimistic within the commission, as work on the document has barely started. But its conclusions will be closely watched outside Brussels — in particular at ESA — as they could signify a new step in the commission's efforts to exert a greater influence over Europe's space activities.

Alison Abbott

Geologist 'victimized' for role in fraud case

New Delhi. A reader in the geology department of the Panjab University in Chandigarh, India, who helped to initiate charges of scientific misconduct against his boss, the geologist Vishwa Jit Gupta, six years ago, claimed last week that he is being victimized by the university. Its senate has blocked a promotion for which he had been recommended by a selection committee.

Arun Ahluwalia says that his promotion was the only one to be blocked by the senate out of 29 approved by the committee. "If this is the reward for exposing frauds in order to preserve scientific values, then only God can save this country," says Ahluwalia, who, together with the Australian geologist John Talent and others, exposed what is widely referred to as the 'Himalayan fossil fraud' perpetrated by Gupta, a senior professor in the geology department at Chandigarh (see *Nature* 338, 613; 1989).

Last year, a commission of inquiry found

Gupta alone — and not his co-workers — guilty of misconduct. But the university senate, which many claim is dominated by political allies and friends of Gupta, handed down what many consider a relatively mild punishment — namely that he should no longer be permitted to teach graduate students or take any administrative responsibility in the university (see *Nature* 371, 368; 1994), while targeting Ahluwalia, the only one of several co-workers who had continued the fight against Gupta.

Last week's decision by the university senate to block his promotion is seen by Ahluwalia as a sign of a corroding value system and "a failure of the Indian scientific community to uphold the cause of scientific integrity". He also says he feels let down by the international scientific community for not applying pressure on either the Indian government, or Panjab University itself.

K. S. Jayaraman

Italy recruits new faculty under old rules

Munich. Italy's minister of research and universities, apparently frustrated at parliamentary delays in approving new rules for the controversial *concorsi* — or national competitions — for university posts, last week announced a new round for 1,200 such positions under the existing rules.

Giorgio Salvini's move comes at a time of increasing public criticism of the *concorsi* system, widely believed to allow academic appointments to be based on personal connections rather than academic merit. The positions to be filled have remained empty for several years as argument has raged about whether appointments should be delayed until new laws aimed at reducing the opportunities for corruption have been passed by parliament.

The move also coincides with the discovery of 50 new cases of alleged corruption within *concorsi* selection committees, made up of professors from appropriate disciplines, which are being considered by the state prosecutor Adelchi D'Ippolito. Unlike the handful of previous challenges to

concorsi decisions, which were heard by administrative courts, the new charges — reflecting the revolt against corruption that has swept Italy over the past three years — are based on an alleged abuse of power, a criminal offence.

Indeed, the current level of public interest in the problems of academic appointments is reflected in the coining by the Italian press of a new term, *cattedropoli* (*cattedra* means chair), which, by analogy with the more familiar word *tangentopoli* used in recent investigations of political corruption, implies corrupt dealing.

Several cases are already coming to trial. Particular attention has been focused on the case of Corrado Manni, an anaesthesiologist at the Policlinico Gemelli in Rome, who treated the Pope when he was shot in St Peter's Square in 1981. Manni chaired one of the *concorso* committees whose activities are being challenged, and has already been denounced by two members of his committee.

All recommendations of *concorsi* com-

mittees must be approved by the National University Council (CUN), a body that advises the minister of universities and research on university issues, as well as by the minister himself. But both are now attempting to disclaim responsibility.

In an article in the daily *Corriere della Sera*, Luigi Frati, dean of the faculty of medicine in Rome's La Sapienza university, argues that the CUN lacks the power to do much more than check that the formal administrative procedures of *concorsi* are carried out. He points out that Salvini had promised a reform of the whole system of academic appointments, but this has not yet happened.

Last week, D'Ippolito called in several high-level 'expert witnesses' to help him to assess the cases now before the courts. These witnesses included Salvini and his predecessor Stefano Podestà.

Asked to explain why he had approved six contested *concorsi* decisions, mostly relating to medical sciences, that were not approved by Podestà, Salvini argued that his duty, like that of the CUN, is limited to approving the way in which the formal procedures of the *concorsi* are carried out, rather than examining individual recommendations for promotion.

By contrast, Podestà told D'Ippolito that he had not signed the six *concorsi* in question because he felt that personal interests might have been operating within the committees. "I believe a minister should not be a paper-pusher," he said.

Earlier this year, Salvini proposed legislation to change the rules of the *concorsi*, allowing more candidates to be selected than the number of positions available (see *Nature* 374, 756; 1995). This would permit individual universities to make their own appointments from a short-list, instead of being assigned a new staff member by a national committee.

Yet despite assurances that the bill would be given priority in parliament, it quickly became mired in committees by discussions and counter-proposals, and is now unlikely to be approved in the near future. As a result, Salvini has decided not to wait for a change in the rules before setting up the next round of *concorsi*, keen to alleviate the block that has built up in the lower ranks of academia appointments.

Even if the bill were to be approved, many believe that it would not go far enough. Fernando Auiti, for example, a professor of immunology at La Sapienza who was also called in as a witness by D'Ippolito, is one of many who argue that objective criteria, such as publication citation rates, should be obligatory within the *concorso* process. "This is normal practice in the world of international science," he points out.

Alison Abbott

El Niño centre is forecasting 'first'

Washington. Scientists and government officials from 40 countries last week endorsed plans for an international research centre to forecast the impact of the El Niño effect on global climate. Supporters of such a centre say that it could eventually generate regular, reliable global forecasts of short-term climate change.

The proposed international research institute will be funded initially by the United States, and located at one of eight competing US universities. It will refine existing models of El Niño and issue global forecasts of its likely impact. Such forecasts will enable farmers to select crops for planting in advance of El Niño's impact — possibly averting huge crop losses in areas where rainfall is influenced by El Niño.

The El Niño–Southern Oscillation, an interaction between the Pacific Ocean and the air above it, is a central factor in short-term climate variation around the world, particularly in the tropics. The warm and cold El Niño cycle has a particularly strong impact on rainfall, and therefore agriculture, in large areas of South America and south and South-East Asia.

The planned centre is seen as an important precedent by climatologists, as it would make available for the first time a full forecasting service based on global climate models. "This is the first step towards long-term climate forecasts," says James Baker, administrator of the US National Oceanic and Atmospheric Administration (NOAA), and co-host of the intergovernmental meeting in Washington last week.

Jack Gibbons, President Bill Clinton's science adviser and the other co-host, told the meeting that early predictions of a drought associated with El Niño in 1992 had enabled farmers in Brazil to maintain production "at near-normal levels", in contrast to a previous, similar pattern in 1987, when production had fallen by 85 per cent.

"The economic benefits of being able to avert crop failures could be worth billions of dollars to us, collectively, every year," said Gibbons. "By working together to improve understanding of the El Niño phenomenon and to produce and disseminate climate forecasts, we can use that knowledge for the betterment of the human condition."

According to NOAA officials, the site of the new international institute will be selected within a few weeks. The agency has budgeted \$4 million in the first year to set up the centre, which could be fully operational within 12 to 18 months.

As well as issuing forecasts and working systematically to refine their accuracy, the institute will seek to speed up the application of climate science to social needs and help poorer countries to make good use of the forecasts.

Last month, the US weather service issued its own forecast that, with sea surface temperatures in the Pacific at or just below normal, the global climate this winter would not be strongly influenced by El Niño conditions. The full forecast is available on the World Wide Web at <http://nic.fb4.noaa.gov>. Colin Macilwain

'Genetic art' builds cryptic bridge between two cultures

Boston. A rare collaboration between about a dozen artists and scientists has led to an exhibition of so-called 'genetic art' at Harvard University in Cambridge, Massachusetts. But one planned exhibit that would have used artificial nucleotide sequences inserted into live *Escherichia coli* bacteria to express an encoded message has fallen foul of the university's safety committee.

The exhibition is the brain-child primarily of Joe Davis, the unofficial and unpaid 'artist-in-residence' in Alexander Rich's Molecular Structures Laboratory at the Massachusetts Institute of Technology (MIT), who says he was inspired by an exchange between the Nobel prize-winners for medicine, Max Delbrück and George Beadle.

In 1958, Delbrück sent a telegram to Beadle at the Nobel prize ceremony in Stockholm containing a string of 229 letters. These formed a coded message in which each of the letters A, B, C, and D corresponded to a DNA base, and a sequence of three bases in turn corresponded to a letter in the English alphabet.

Beadle successfully deciphered the message "break this code or give back the Nobel prize", and responded with a coded message of his own. Delbrück replied with yet another message, this one in the form of a set of coloured toothpicks which, when decoded, produced the message: "I am the riddle of life. Know me and you will know yourself."

Davis has now used PCR technology to take the concept a step further. Last year, he teamed up with biologists at MIT and the Free University of Berlin to synthesize DNA molecules composed of a sequence of 174 bases. The bases are arranged in triplets, each standing for a separate letter which together spell out Delbrück's "riddle of life" message.

The Harvard show consists of various sculptures — including a helical array of broomsticks, a string of coloured toothpicks, and a rack of test tubes containing light-responsive pigments — each illustrating this message in a different way.

Davis argues that the exhibition is a reflection on the idea that, despite what has been discovered about it, DNA still remains "one of the principal riddles of life". The

sculptures, he adds, are about "making the invisible visible — an effort common to art and science". But although his efforts have been appreciated on artistic grounds, their biological aspects have been controversial.

In September, Davis was given written permission from Harvard's Committee on Microbiological Safety (COMS) to display the recombinant *E. coli* organisms containing the "riddle of life" DNA at the exhibi-

tion, provided that he adhered to certain conditions.

Davis says he interpreted the conditions as allowing him to display "viable cultures" provided they were properly sealed and refrigerated, and kept secure. But the committee, in a subsequent cla-

rification of its position, told Davis that the cultures had to be "inactivated" with formaldehyde and chloroform.

John Mekalanos, the chair of the medical school's safety committee, says the committee made "perfectly reasonable suggestions that should not infringe on [Davis's] artistic freedom". But Davis says he was reluctant to "kill" organisms he considers to be "harmless", and has instead chosen to display the synthetic DNA itself in two leakproof containers in a refrigerator — steps that would meet the committee's stipulations.

Sarah Zehr, a biological anthropologist at Harvard who helped to set up the show, says that she agrees with Davis that "fixing" the cultures was unnecessary, as "the DNA is inert and doesn't code for anything". But she accepts that recombinant DNA is still a politically charged issue in Cambridge, site of a fierce conflict between Harvard and the city's Mayor Al Vellucci in the late 1970s. "There are still some old sentiments here which die hard," she says.

Zehr claims that one of the benefits of the show is that "it brings artists and biologists together, two groups which normally have little to do with each other". Mekalanos says that he is also intrigued by the idea of using DNA as "a wholly new form of communication and artistic expression". He adds: "we're giving [Davis] a chance to display these biological materials because we feel it could be historically significant." The exhibition remains open until the end of the month.

Steve Nadis

Democrats rally to defend science from Congressional cuts

Washington. Democrat members of the US Congress are mounting a late bid to put proposed Republican cuts to research and development on the public agenda, just as the budget battle between President Bill Clinton and the Republican majority in Congress reaches its climax.

Last week, six Democrat senators joined Clinton administration officials to lambast what one of them — John Glenn of Ohio — termed the "myopic, Luddite view" of the Republicans on supporting research. "These modern-day Luddites would cut back on basic, non-defense research and development by 30 per cent by the year 2002," said Glenn, a former astronaut.

His views were echoed by John Rockefeller (Democrat, West Virginia), who said that the public did not understand the impact of the planned cuts. "This is as dangerous as [planned cuts in] Medicare, because it affects jobs for our young people in the future," he said.

Meanwhile, in the House of Representatives, George Brown (Democrat, California), the senior Democrat on the science committee, has set up a 16-member research and development (R&D) task-force, which will seek to raise public awareness of the difference between Democrat and Republican positions on R&D.

Brown is worried that scientists — particularly in universities — have accepted Republican assurances that basic science will be protected, and hopes to shake them out of any such complacency. The task force, which includes several Democrat members of the House appropriations committee, plans to meet Clinton soon to discuss his budget plans for the 1997 financial year (which begins on 1 October 1996), which is due in February.

The Clinton administration's Council of Economic Advisers backed up the Democrat Senators' attack by releasing a short report on the role of federally funded R&D in boosting the US economy. This predicts that industry-funded R&D will drop if government-funded R&D does so.

Challenged with National Science Foundation data showing that the Clinton administration has already cut R&D spending (see *Nature* 378, 3; 1995), Laura Tyson, Clinton's economics adviser, said that "our budget tries to maintain R&D spending, as a proportion of the total budget".

Earlier opposition to Republican budget plans has been undermined by bad relations between Clinton and Democrats in Congress. But Democrats have since closed ranks in an effort to land Republicans with the blame for the long-awaited budget 'train-wreck'.

Colin Macilwain.

IMAGE
UNAMIGABLE
FOR A COPYRIGHT
FOR A SHORT
REASONS

Art meets science? Coloured broomsticks mark out the sequence of DNA bases, echoing Delbrück's message.

French nuclear tests at Mururoa

SIR — In connection with your news report about French nuclear tests at Mururoa Atoll (*Nature* 377, 91–92; 1995), I should like to clarify the Greenpeace assertion about traces of caesium-134 (“a known product of nuclear fission”) indicating leakage of radioactivity around the atoll from past French underground nuclear tests. This is very doubtful. Caesium-134 is not a direct product of nuclear fission, but is created from caesium-133 by the absorption of a neutron. As a consequence, caesium-134 is not significantly produced in a nuclear explosion, but is produced in prodigious quantities in nuclear reactors. (See Table 1.3 of Radioactivity after Chernobyl SCOPE 50, edited by Warner and Harrison, in which radioactive releases from nuclear detonations and nuclear reactor accidents are compared.) Indeed, one of the first tests undertaken of a radioactive plume is to examine it for the presence of caesium-134 to establish if the source was a weapon or a reactor. This was done, for example, by the Swedes in 1986 when they were the first in the West to announce the reactor accident at Chernobyl.

In a French study (Bourlat and Martin *J. env. radioactivity* 17, 13–29; 1992) to verify if caesium-134 traces were present in Mururoa lagoon as reported by the Cousteau Foundation in 1987, no trace was found. Measurements of radionuclides in seawater and plankton collected outside Mururoa Atoll were also conducted under the sponsorship of the International Atomic Energy Agency (IAEA) by three independent international laboratories, to provide an inter-comparison of results (IAEA/AL/044, (AEA-ILMR Report 48, July 1991). The only man-made radionuclides they detected were strontium-90, caesium-137 and plutonium-238, 239 and 240.

The question remains: where did the caesium-134 detected by the Cousteau Foundation come from? Bourlat and Martin refer to a New Zealand study that found evidence that it could have come from laboratory samples contaminated with Chernobyl radioactivity (Cambray *et al. New Zealand Radioactivity Annual Reports* 1986–1990).

For the past two years, I have served as the executive director of an international collaborative study of the radioactivity from nuclear tests, called RADTEST. Sponsored by SCOPE (the Scientific Committee on Problems of the Environment, based in Paris), our study has focused on human exposure and the effects from nuclear weapons test fallout. We have found that the principal hazard is from local early fallout from atmospheric ground or near-ground bursts. A lower level impact is from global delayed fallout

from atmospheric ground and air-bursts. There is no significant human exposure from underground bursts except when the radioactivity vents from the surface, which has occurred on numerous occasions. Limited studies of migration from contained underground tests in Nevada in the United States and at Semipalatinsk in Russia indicate very slow movement, and hence only a potential human hazard in the very long term. If the proposed French tests are properly contained, there should be minimal human exposure. As to possible environmental damage to the atoll, this is beyond the scope of the RADTEST study.

Charles S. Shapiro

*Regional Atmospheric Sciences Division,
Lawrence Livermore National Laboratory,
PO Box 808, L-262,
Livermore, California 94551, USA*

SIR — Your leading article “France and nuclear illusion” (*Nature* 377, 89; 1995) smacks of typical Anglo-Saxon solidarity and double standards.

“France has no better case than Britain for maintaining separate nuclear forces”, you state. OK. But then what is the case for Russia or for the United States? For the latter because they (and only they in history) have already used it?

“Accidents of postcolonialism should have given France the right...”. What do you think Australians and New Zealanders are? Natives? Of course, in these countries there are few natives left who could protest about anything!

J. E. Dumont

*Université Libre de Bruxelles,
Campus Hôpital Erasme,
Route de Lennik 808,
1070 Brussels, Belgium*

Historic apoptosis

SIR — In his interesting film review, Peter Tallack¹ asks “Why... has something so fundamental as apoptosis been ignored for so long?” The popular view, reiterated in countless papers, is that apoptosis was missed until unexpectedly brought to light in 1949 or 1972 or still more recently.

The truth is that naturally occurring cell death (including the concept of apoptosis, though not the name) was a flourishing subject at the end of the nineteenth century, and was systematically reviewed by Barfurth each year between 1893 and 1913 in the *Ergebnisse der Anatomie und Entwicklungsgeschichte* under the heading “Involution”, but then went out of fashion.

The first report that cells die naturally in development was Carl Vogt's account² of dying notochordal and cartilaginous cells in 1842, which was remarkably soon

after the establishment of the cell theory by Schleiden and Schwann between 1839 and 1842. The next landmark was August Weismann's detailed description³ in 1864 of massive cell death in pupating diptera, which triggered others into writing a dozen papers on this subject between 1870 and 1890, although Weismann himself gave it up because of an eye disease and achieved immortality theorizing about the immortal germ plasm.

Other landmarks were Stieda's description in 1872 of chondrocyte death during endochondral ossification, Metschnikoff's account in 1883 of phagocytosis associated with cell death in the muscles of metamorphic toads, and Beard's discovery in 1889 that an entire population of neurons is lost in fish embryos. The first report of scattered cell death in a tissue destined to remain was that of Felix⁴ in 1889 on developing mammalian muscle. Neuronal death among developing motoneurons and spinal ganglion cells was described in 1906 by Collin⁵, who understood the numerical importance of the phenomenon, and argued that an overproduction of cells followed by degeneration of the excess may be common to all tissues.

The occurrence of different kinds of degeneration was common knowledge to pathologists by the 1860s, but the first clear morphological description of apoptosis was that of Flemming⁶ in 1885 in a study of naturally regressing ovarian follicles. He called it “chromatolysis”, and the term was widely used for this kind of cell death during the next 30 years. As has recently been pointed out⁷, the occurrence of this type of cell death in breast cancer was reported in 1892, and the need of it to counterbalance mitosis was argued by Gräper in 1914.

Most of the nineteenth-century reports were written in German, and the reduced interest in cell death after about 1914 seems to have been due partly to the increasing proportion of scientists unable to read that language. The torch was, however, kept alight by embryologists such as M. Ernst, H. Fell and A. Glücksmann in the period between the two wars, followed by V. Hamburger and R. Levi-Montalcini over the subsequent two decades. The rest is common knowledge.

Peter G. H. Clarke

*Institute of Anatomy,
Stephanie Clarke
Institute of Physiology,
University of Lausanne,
1005 Lausanne,
Switzerland*

1. Tallack, P. *Nature* 377, 393 (1995).
2. Vogt, C. *Untersuchungen über die Entwicklungsgeschichte der Geburtshelferkroete (Alytes obstetricans)* (Jent und Gassmann, Solothurn, 1842).
3. Weismann, A. Z. f. wissenschaft. Zool. 14, 187–336 (1864).
4. Felix, W. Z. f. wissenschaft. Zool. 48, 224–259 (1889).
5. Collin, R. *Névrase* 8, 181–309 (1906).
6. Flemming, W. *Arch. f. Anat. und Physiol. Jahrgang 1885*, 221–244 (1885).
7. Majno, G. & Joris, I. *Am. J. Path.* 146, 3–15 (1995).

Colloidal crystals model real world

Neat Australian experiments have shown that an artificial three-dimensional colloid can be a good model of the real world of crystallization in the lab as well as on paper.

EVERYBODY would like to know more about two of the most common phenomena in physical science — the melting of solids and the inverse process, their crystallization. Of course, the gross features of each process are easily described. A pure material melts at a particular temperature, called the melting point. There is a positive latent heat and, given information about the compressibility of the liquid and solid phases, it is possible to calculate the change of the melting point with pressure. Simple thermodynamics, that is, first done 150 years ago by William Thomson (later Lord Kelvin).

The trouble is that this stark statement cannot apply to what happens on a microscopic scale. Improperly to put the difficulty in anthropomorphic terms, how does a whole crystal being warmed *know* that the point is approaching when each of its atoms or molecules is to forsake imprisonment on the crystal lattice for the freedom to move about at will? Luckily, there is a means of communication within all crystals: the lattice vibrations which, quantized by the procedures worked out by Max Born and others in the 1930s, each involve every atom in the crystal. But if that is how the warning is delivered, how can inanimate atoms and molecules remember what it is?

Part of the answer no doubt lies in phrases such as the 'roughening transition', the observation that physical changes in the surface of a solid crystal appear to show up in advance of melting in the strict sense. It looks as if some parts of the crystal are *preparing* (more anthropomorphism) to break away. But dutifully they *wait* until the whole crystal is ready, and only then they go free. It would be an interesting exercise to calculate what blend of photons sets this process in train.

On the face of things, crystallization should be more easily understood, but that is not what closer inspection shows. Crystals cannot appear as entire objects, but must grow, for one thing. That requires a register for the regularity of their lattices, which must be a nucleation centre. That is a microcrystal of some kind, perhaps just a cluster of a few atoms or molecules formed stochastically in the melt, perhaps a crystal of some more refractory material that happens to have similar lattice constants or perhaps just a roughness of appropriate microscopic dimensions on the boundary walls of the container. Another difficulty (but really it is support

for the doctrine of nucleation) is that supercooling of a liquid below the melting point is commonplace.

Over the years, the concept of nucleation has been recognized as a prerequisite of crystallization. The trouble here is that it has hardly ever been observed. Especially if there is supercooling, a crystal will form from a melt too quickly to be followed with all the high-speed cameras in the world (but people keep trying). And from time to time, scraps of new information do nevertheless come to light.

A band of brothers at the Royal Melbourne College of Technology has meanwhile, over at least a decade, been cultivating a model of the process of crystallization that stands in relation to the real thing as does, say, the behaviour of fluid vortices in a rotating drum to the evolution of the weather in the atmosphere of the spinning Earth. Their starting point is the notion that the slow formation of crystalline aggregates in a suspension of colloidal particles is a representation of what happens when melts are cooled. The difference is that the speed with which ordered structures form in colloidal suspensions is slow enough to be followed at leisure.

First, then, make your colloid. W. van Megen, apparently the leader of this enterprise, settled some years ago for colloids made from polymethylmethacrylate spheres (400 nm in diameter, or no bigger than a wavelength of red light), coating them with a 10 nm layer of covalently bound macromolecules to ensure repulsive forces between neighbouring spheres and good interaction with the solvent.

In earlier reports (P. N. Pusey & W. van Megen *Nature* **320**, 340–342; 1986 and W. van Megen & S. M. Underwood *Nature* **362**, 616–618; 1993), the group had effectively established that the analogy between the formation of ordered structures of these spheres and the process of crystallization in real molten solids is more or less exact. Now they have gone a step further, and have shown not merely that nucleation is an essential feature of crystallization, but that nuclei form more quickly as crystallization proceeds (J. L. Harland, S. J. Henderson, S. M. Underwood & W. van Megen *Phys. Rev. Lett.* **75**, 3572–3575; 6 November 1995).

What happens with a colloid of this kind depends crucially on the volume fraction of the colloid particles, in this case the polymethylmethacrylate spheres. If the fraction

exceeds 0.58, the colloid forms a glass: the spherical particles cannot diffuse around each other. In a more dilute suspension with a volume fraction less (but not too much less) than 0.494, on the other hand, the particles form a regular array as if they were part of a crystal; that is the freezing point. The melting point (but remember that the variable is not temperature but volume fraction or concentration) lies above the freezing point and below the glass transition.

The interest of what follows is that it is possible actually to carry out measurements. Indeed, the trick is to use the scattering of laser light from a monochromatic laser as if it were an X-ray beam used in Bragg diffraction mode. The system has then to be calibrated by recording the signal from incoherently scattered light with an array of photodetectors. A cubical glass box one centimetre in all directions makes the working laboratory space from which light is scattered and otherwise collected. Because the colloidal particles hold themselves in crystalline arrays only tenuously, it is child's play to reduce crystal to its constituent polymethylmethacrylate spheres; simply arrange to have the 1-cm cube box turned end over end for a day or so, whereupon whatever it may contain will fall apart under shear stress.

Unpromising though the analogy between an ordered colloidal structure and a crystal may be, the outcome of the measurements is encouraging. Much as in the real world of molecular crystals, colloidal crystals form at such a pace that the nucleation rate appears to increase as the volume not occupied by ordered colloidal particles shrinks. That appears to accord with what has been calculated about the properties of plastic spheres in a solvent with which they interact.

What is to be learned from studies of this kind? A little and a lot. The simplest truth is that nucleation does indeed play a crucial role in crystallization. It is also unsurprising that, with the passage of time, the volume density of ordered patches of crystallinity at first grows then shrinks; crystals are created at the nuclei formed in increasing numbers, but then also begin to swallow up each other. But what is hard to understand is that the rate at which crystallization nuclei appear seems to increase as the material from which they may be made shrinks in volume. As yet, there is only hand-waving to sustain that point.

John Maddox

Knee-deep in cosmic rays

Stephen P. Reynolds

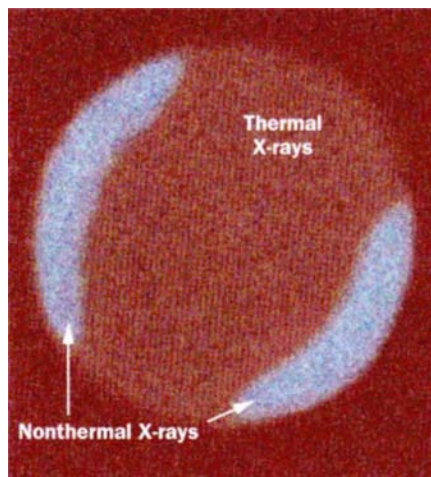
COSMIC rays — the electrons and ions of relativistic velocities that bombard the Earth from all directions — were first observed in experiments with balloon-borne electroscopes over 80 years ago. The problem of their origin has attracted some of the great minds of twentieth-century physics, most famously Enrico Fermi¹. The general view² is that the vast bulk of cosmic-ray ions and electrons, up to energies of about 10^{15} eV (1,000 TeV), are accelerated in the shock waves of supernova remnants, in a process that is a descendant of Fermi's original idea. But direct evidence for particles of these enormous energies in supernova remnants has been lacking. Now Koyama *et al.*³, as they describe on page 255 of this issue, have evidence for the presence of electrons with energies over 100 TeV in at least one such remnant, that of the supernova known from historical records in AD 1006.

Cosmic rays are detected at the Earth by satellites, balloons and ground-based air-shower detectors. Their energies are observed to range from below 1 GeV, where they disappear into the solar wind, up to the astonishing energy of over 10^{20} eV, or around 10 joules. The highest-energy events are detected at the Earth's surface as huge sprays of debris resulting from collisions by unknown particles with air molecules at the top of the atmosphere. The spectrum is close to a power law in energy over most of this range, with an index of about -2.7 below 1,000 TeV, where a slight steepening occurs (the 'knee') to an index of about -3.1 . Cosmic-ray electrons are observed up to about 1,000 GeV. Above this energy, the electron flux, already lower than that of ions by about a factor of a hundred, drops off more steeply and becomes lost in a background of proton-induced events.

It is generally thought that Galactic sources, most likely interstellar shock waves from supernovae, produce the cosmic rays up to the 'knee', whereas some other process, perhaps extragalactic, is responsible for the more energetic component. (Pulsars do produce extremely energetic electrons, but are not thought likely to account for a significant fraction of Galactic cosmic rays.) In the process of diffusive (Fermi) shock acceleration², energetic particles move upstream of the shock, scatter from magnetic turbulence convected in with the incoming flow and acquire extra energy. Some of them return downstream and are again reflected from downstream turbulence, again gaining energy. The cycle can repeat indefinitely, until particles are lost down-

stream or reach energies at which the required resonant scattering waves are absent.

For cosmic rays below the 'knee', supernova remnants make handy culprits; they result from highly energetic (10^{44} joules) explosions which are regularly observed in galaxies like our own (although none has been observed in the Milky Way since Kepler's supernova of 1604). Their shock waves expand into the interstellar medium for up to 100,000 years, reaching sizes of tens to hundreds of light years. Supernova remnants are associated with strong radio synchrotron



Outline of the X-ray morphology of supernova remnant 1006. The brightest regions, and those now shown to have nonthermal spectra, are the opposing blue arcs, whereas the remainder of the remnant shows fainter, less energetic thermal emission.

emission from relativistic electrons, providing direct evidence for the production of electrons with energies at least into the 1–10 GeV range. But it is a large extrapolation from this evidence to the conclusion that the remnants can provide the entire Galactic component of all cosmic rays — it's a long way from 10 GeV to 1,000 TeV. There is as yet no firm evidence for the acceleration of ions in supernova shock waves, and an early estimate⁴ of the maximum energies that could be produced by shock acceleration suggested that the 'knee' could not in fact be reached.

In addition to being potential cosmic-ray factories, supernova remnants are bright thermal X-ray sources, their spectra revealing the elemental abundances of the stars that produced them. Whilst studying one such object, the galactic remnant of the supernova of AD 1006, Koyama *et al.*³ have found a convincing resolution to a long-standing puzzle. Unlike other remnants of historical supernovae, this had been found to have an X-ray spectrum

devoid of individual spectral lines, with a continuum shape much better represented by a power law than by the expected exponential cut-off from thermal bremsstrahlung. Although early suggestions were made that the X-rays might be synchrotron emission from highly energetic electrons⁵, the discovery of weak oxygen lines below 1 keV seemed to favour a thermal origin, and complex thermal models were produced⁶ which managed to suppress line emission.

Koyama *et al.* find, however, that the power-law spectrum is alive and well in the bright opposing limbs of the object, shown in blue (more energetic photons) in the figure. Their observations, made with the Japanese ASCA X-ray satellite, show that spectra from the fainter areas of the limbs, and from the centre, are normal and line-dominated. Koyama *et al.* conclude that the thermal model cannot explain the emission from the bright areas, and that the most likely origin here is synchrotron emission from electrons. The emitted photon energy is linked to the electrons' energy by the magnetic field strength, estimated to be 10^{-5} gauss, requiring that the radiating electrons range in energy to 100 TeV and above — within a factor of 10 of the 'knee'.

Confirmation of the nonthermal nature of the X-ray emission is a large step towards the confirmation of supernova remnants as the source of Galactic cosmic rays below 1,000 TeV. The putative process of shock acceleration operates differently for electrons and ions of low energies, but at higher energies such that both ions and electrons are relativistic, the processes by which they produce, and scatter from, magnetohydrodynamic waves are identical. Thus the evidence that electrons find the appropriate conditions to attain energies over 100 TeV strongly suggests that those conditions are accessible to ions as well.

Confirmation of the nonthermal nature of the emission from supernova 1006 could be made when we have observations at higher energies, such as should be possible from the X-ray Timing Explorer scheduled to be launched this month. But already, the direct demonstration that supernova remnants have the wherewithal to produce 100 TeV particles has gone far towards solving at least part of the mystery of cosmic rays. □

Stephen P. Reynolds is in the Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA.

1. Fermi, E. *Phys. Rev.* **75**, 1169–1174 (1949).
2. Blandford, R. D. & Eichler, D. *Phys. Rep.* **154**, 1–75 (1987).
3. Koyama, K. *et al. Nature* **378**, 255–258 (1995).
4. Lagage, P. O. & Cesarsky, C. J. *Astr. Astrophys.* **125**, 249–257 (1983).
5. Reynolds, S. P. & Chevalier, R. A. *Astrophys. J.* **245**, 912–919 (1981).
6. Hamilton, A. J. S., Sarazin, C. L. & Szymkowiak, A. E. *Astrophys. J.* **300**, 698–712 (1986).

From stress to cognition

Stafford L. Lightman

IN response to acute physical or psychological stress, the hypothalamus, pituitary and adrenal glands are rapidly activated; the result is an increase in circulating glucocorticoid hormones, which can be life-sustaining, particularly because of their metabolic effect on elevating blood glucose levels. Two papers in this issue further our understanding of the role of corticotropin-releasing factor (CRF), a 41-amino-acid neuropeptide found in the stress responsive cells of the hypothalamus which initiate the whole sequence of endocrine responses. Vaughan *et al.*¹ have discovered another member of the CRF family in mammals, while Behan *et al.*² have studied the connection between lowered levels of CRF and Alzheimer's disease, providing a new perspective on possible strategies for treating the disease.

Corticotropin-releasing factor is secreted by the hypothalamus in response to stress, and through hypothalamus-pituitary portal blood is targeted to the corticotrope cells of the anterior pituitary. These cells are in turn stimulated to release the hormone corticotropin (or adrenocorticotrophic hormone, ACTH), which binds to receptors in the adrenal cortex and activates the release of glucocorticoid hormones. Before 1981 there was a great deal of indirect evidence for CRF's existence, but it was only then that CRF-41 was cloned³ and we could at last sigh with relief that we had found what had been the missing link in the regulation of the hypothalamic-pituitary-adrenal stress response. Before long, however, chinks appeared in our confidence that CRF was simply a hypothalamic neurohormone which acted directly on anterior pituitary corticotropes. The new papers^{1,2} help to explain the gradual accumulation of confusing data.

The first surprising finding was that CRF appears in high concentrations in peripheral plasma during pregnancy^{4,5}; and although it reaches similar levels to those in the hypothalamic-pituitary portal blood of stressed experimental animals, plasma ACTH levels remain normal. The subsequent discovery and cloning of a CRF-binding protein in human plasma⁶ explained the dissociation between plasma CRF and ACTH levels because the binding protein can inactivate CRF. But the roles of circulating CRF and of the CRF-binding protein — also found in the central nervous system (CNS) — remained opaque.

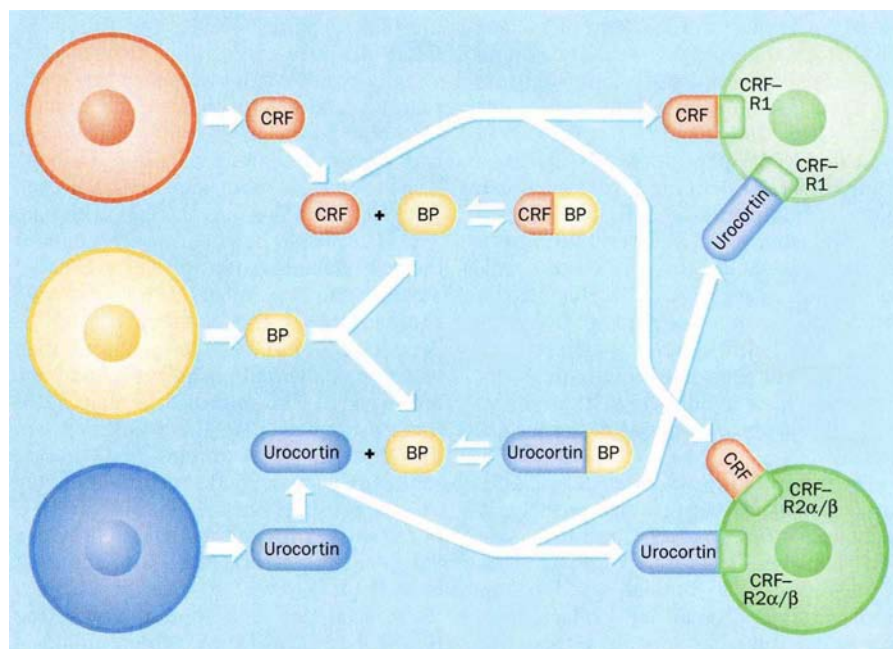
The next area of confusion concerned the sites of synthesis and function of peripheral CRF, or of similar peptides with CRF-like immunoreactivity. Given

the high levels of CRF during pregnancy, it was not too surprising that high levels of CRF-like peptide⁷ and CRF messenger RNA⁸ were found in human placenta and probably have local biological effects. Further reports, however, also began to appear suggesting that CRF or CRF-like compounds could be found in other peripheral sites and could have local effects on inflammation⁹⁻¹³.

The precise nature of these peripheral CRF-like peptides was uncertain. There were, however, several lines of evidence that hinted that the binding protein might

ian peptides urotensin and sauvagine, which share over 50% of their residues with CRF.

With this as background, Vaughan *et al.* (page 287 of this issue¹) have localized urotensin-like immunoreactivity in rat brain using polyclonal antisera, and have constructed a complementary DNA library from an area of high immunoreactivity (the rat midbrain) and screened it with a urotensin probe. The upshot is the cloning of a cDNA encoding a peptide with 63% homology to suckers urotensin and 45% to rat CRF. This peptide, which the authors call urocortin, is slightly less effective than CRF in inducing cyclic AMP in cells transiently transfected with CRF-R1, but is 10 times as potent in cells expressing CRF-R2β. The



Interactions between corticotropin-releasing factor (CRF), urocortin, CRF-binding protein (BP) and the CRF receptors (CRF-R1, CRF-R2α and CRF-R2β). There may also be competition between urocortin and CRF for binding to BP, and other undefined CRF-like peptides may compete for BP and CRF receptors.

interact with alternative CRF-like ligands — the high levels of CRF-like peptides with only very low levels of CRF mRNA in synovial tissue; the fact that CRF-binding protein in the brain occurs in many sites that have no measurable CRF; and the presence of considerable levels of binding protein in human males and non-pregnant females who do not have a placental source of CRF. Data on CRF receptors also pointed to the existence of alternative ligands. Two different genes were described encoding CRF-R1 and CRF-R2 receptors, there being two different splice variants, α and β, for CRF-R2 (refs 14, 15). CRF-R2 was found to have a wide distribution with obvious mismatch to the known sites of CRF synthesis. The CRF-R2 receptors were also less sensitive to mammalian CRF than to activation by the fish and amphib-

peptide is a potent activator of ACTH secretion through CRF-R1 receptors both *in vivo* and *in vitro*, although there is no obvious source of urocortin which would have access to portal blood. Urocortin also has hypotensive effects when given intravenously, and from immunoreactivity data it appears to occur peripherally as well as in the CNS. Indeed, many of the cardiovascular and immunological effects of exogenous CRF may be a reflection of the normal actions of endogenous urocortin, although the relative importance of CRF and urocortin in CNS-mediated responses to exogenous CRF needs to be defined.

Urocortin also has a high affinity for binding to the CRF-binding protein, and this links in with the studies of Behan *et al.* (page 284 of this issue²). These authors investigated possible interactions of CRF

and its receptors and binding proteins in patients with Alzheimer's disease. They focused on the low cerebrospinal fluid concentrations of CRF in such patients and the markedly reduced CRF peptide concentration and increased CRF receptor expression in Alzheimer's brain tissue. They studied 10 brains from Alzheimer's patients and 10 controls, *post mortem*, confirming the reduction in cortical CRF immunoreactivity but finding no changes in CRF-binding protein immunoreactivity. Whereas 40% of cortical CRF was bound to binding protein in normal brain, more than 60% was bound in Alzheimer's brain. This suggests a decreased availability of the already reduced levels of CRF in patients with the disease.

Because CRF has known cognitive-enhancing effects in rats, Behan *et al.* tested whether agents that displace CRF from the complex between it and its binding protein had similar effects to direct administration of exogenous CRF. CRF and human CRF 9-33 (a truncated CRF peptide which has high affinity for CRF-binding protein and can displace endogenous CRF, but has no intrinsic activity at CRF receptors) both had memory-enhancing effects on rats, but only the exogenous CRF had anxiogenic (anxiety-arousing) side-effects in a water-maze behavioural test. These side-effects would limit the potential use of CRF analogues in human patients, so this alternative approach might be important clinically.

The interpretation of these data is difficult, particularly because the changes in Alzheimer's brain may be due to loss of CRF neurons, decreased CRF synthesis or increased CRF breakdown. Furthermore, CRF 9-33 could be displacing not only CRF but also urocortin from the binding protein. But if free CRF is involved in cognitive behaviour in humans, it does open a potential therapeutic approach with agents that can dis-

sociate CRF and urocortin from their complex with CRF-binding protein. Certainly, with the new information we have it will be important to unravel the possible cognitive effects of urocortin and the dynamics of the interplay between urocortin, CRF and CRF-binding protein in the CNS.

There are almost certainly more twists to come in the CRF family story. The emergence of a second CRF agonist is

likely to herald the discovery of other CRF-like peptides, and it will be some time before we can feel comfortable that we really understand the function of CRF-like peptides and their regulation both in the CNS and the rest of the body. □

Stafford L. Lightman is in the Department of Medicine, University of Bristol, Bristol Royal Infirmary, Marlborough Street, Bristol BS2 8HW, UK.

ICE CORES

Profiles of the past

David A. Peel

It is now more than two years since the United States' GISP2 and European GRIP deep ice-core drilling teams, operating in parallel from camps 30 km apart, reached bedrock at Summit in central Greenland. Since then, isotopic, chemical and physical analysis of more than 40 components in the ice has delivered fascinating insights into changes in climate and in atmospheric chemistry throughout the last glacial to interglacial cycle¹. At a meeting in September*, it was clear that comparison of the Summit records is benefiting from new high-resolution data on two globally well-mixed gases, methane and oxygen. The approach is helping to forge a long-awaited direct link between the Greenland, Antarctic and marine sediment palaeoclimate records, opening up the prospect for a truly bipolar view of the world's climate. It has also provided convincing evidence that the apparent instabilities in the climate of the last interglacial period before our own (the Eemian, 125 to 115 kyr ago) revealed in the GRIP core² are probably artefacts.

Participants at the joint meeting of the two ice-core communities were able to appreciate the benefits of the duplicated drilling operations at Summit, sometimes considered an expensive luxury. This was the first time that an attempt in Greenland had been made to reach ice deposited during the last interglacial at a level sufficiently far above the bedrock where there were good prospects to recover a stratigraphically secure, high-resolution record of this distant analogue of our present climate.

In the event, the two records do agree remarkably closely back to 110,000 years ago³, down to a point some 250 m above bedrock, and confirm beyond doubt the existence of a series of remarkable rapid climate oscillations during the last glacial, which ended about 11 kyr ago. These features had been observed previously only

much closer to bedrock, where potential perturbations due to flow disturbances could have occurred. But below this depth, and extending into the ice corresponding to the last interglacial, the records diverge, and structural disturbances in the ice first become apparent. Detailed comparative studies of these features in the two cores have since been made, but until now have proved inconclusive⁴.

High-resolution measurements of two atmospheric gases, CH₄ (J. Chappellaz, LGGE, Grenoble and Univ. Berne) and $\delta^{18}\text{O}$ in air trapped in the ice (B. Malaize and M. Bender, Univ. Rhode Island), presented at the meeting, have allowed convincing stratigraphic linking of records from the GISP2, GRIP and Vostok, Antarctica, cores back to 150 kyr ago. Methane is well mixed throughout the troposphere, and with the exception of small interhemispheric concentration differences (themselves interesting in relation to the long-term controls on CH₄ in the atmosphere) the longer-period fluctuations are clearly global and can be identified and matched in all the cores. Similarly, $\delta^{18}\text{O}(\text{air})$ is believed to be mainly controlled by the oxygen isotope content of sea water, which is in turn determined by the volume of ice locked in the polar ice sheets. It therefore constitutes an independent global stratigraphic marker.

With this much firmer frame of reference it is now possible to embark more confidently on debate on the more subtle interhemispheric differences observed in, for example, CH₄ in different stages of the last glacial cycle, and to examine the relative phasing of climate events in the two polar regions, including the open question of which hemisphere led the warming at the end of the last glaciation. This will be key evidence in exploring how the climates of the Northern and Southern Hemispheres are connected under widely different climate conditions and over different timescales. What, for instance, is the relative importance of

1. Vaughan, J. *et al.* *Nature* **378**, 287–292 (1995).
2. Behan, D. P. *et al.* *Nature* **378**, 284–287 (1995).
3. Vale, W., Speiss, J., Rivier, C. & Rivier, J. *Science* **213**, 1394–1397 (1981).
4. Sasaki, A. *et al.* *J. clin. Endocr. Metab.* **59**, 812–814 (1984).
5. Suda, T. *et al.* *J. clin. Invest.* **76**, 2026–2029 (1985).
6. Potter, E. *et al.* *Nature* **349**, 423–426 (1991).
7. Sasaki, A. *et al.* *J. clin. Endocr. Metab.* **67**, 768–773 (1988).
8. Grino, M., Chrousos, G. P. & Margioris, A. S. *Biochem. biophys. Res. Commun.* **148**, 1208–1214 (1987).
9. Webster, E. L. & De Souza, E. B. *Endocrinology* **122**, 609–617 (1988).
10. Stephanou, A., Jessop, D. S., Knight, R. A. & Lightman, S. L. *Brain Behav. Immun.* **4**, 67–73 (1990).
11. Karalis, K. *et al.* *Science* **254**, 421–423 (1991).
12. Crofford, L. J. *et al.* *J. Immun.* **151**, 1587–1596 (1993).
13. Wei, E. T., Gao, G. C. & Thomas, H. A. in *Corticotropin-Releasing Factor Ciba Fdn Symp.* 172 (eds Chadwick, D. J. *et al.*) 258–268 (Wiley, Chichester, 1993).
14. Chen, R., Lewis, K. A., Perrin, M. H. & Vale, W. W. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8967–8971 (1993).
15. Lovenberg, T. W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **92**, 836–840 (1995).

*GISP2-GRIP Joint Workshop, Wolfeboro, New Hampshire, 16–21 September 1995. Proceedings to be published as a joint special issue of *J. geophys. Res. Oceans and Atmospheres*.

atmospheric processes, such as greenhouse gas forcing, and deep ocean circulation? Prospects of making the long-awaited connection between the ice-core records and the records from deep-sea sediments came much closer when the new high-resolution data on $\delta^{18}\text{O}(\text{air})$ from Summit and Vostok were shown to correlate closely both with the $\delta^{18}\text{O}$ deep-sea record and with the summer insolation at 65°N through the past 100 kyr (Malaize and Bender).

For ice over 110 kyr old, the new bipolar data for the gases, together with more evidence from joint detailed comparisons of the structural properties of the cores, now make an overwhelming case that both cores have suffered stratigraphic disturbance. Therefore, what had appeared to be sudden cold events in the Eemian seen in the GRIP core² are most probably artefacts caused by flow disturbances. It had been thought that these unexpected features might indicate that the relative stability of our climate over the past several thousand years is abnormal, even for warm parts of the glacial cycle. The rather worrying implications prompted a search of other high-resolution records of this important distant analogue of modern climate, but produced no persuasive corroboration.

Participants were left with grounds for optimism that by using a fingerprint based on these two constituents it should prove possible at least to identify and characterize the genuine Eemian ice in the Greenland ice, with some prospect ultimately of reconstructing an ordered sequence. The latest physical modelling (E. D. Waddington, Univ. Washington), which is benefiting from higher-resolution radar soundings of the layering deep in the ice sheet (H. Miller, Alfred-Wegener Inst.), is starting to identify some of the factors that may have disturbed the ice much further away from bedrock than hitherto encountered. It is possible, for instance, that the lateral migration of the ice divide in central Greenland during the last glacial-interglacial cycle was greater than we had thought. And substantial rheological differences have been discovered between glacial and interglacial ice that could promote folding or boudinage (D. Dahl-Jensen, Univ. Copenhagen).

Part of the meeting was devoted to a re-examination of some of the fundamentals of ice-core science. There has been a recent upsurge in debate on one of the primary climate proxies, the stable isotope content of the ice, and especially on the respective roles of local temperature or changing origin of precipitation in shaping these records⁵. To infer temperature shifts, it has been broadly assumed that present-day relationships between δD or $\delta^{18}\text{O}(\text{air})$ and local temperature in different regions are valid also for interpreting temporal changes. Although model com-

parisons indicate that the relationship may also be valid under glacial conditions (J. Jouzel, LGGE, Grenoble), we know that the Greenland climate and atmospheric circulation are sensitive to shifts in the position and strength of the polar frontal system, so considerable effort has been put into independent validation of the stable isotope thermometer for Summit.

Present-day profiles of temperature through the ice sheet, which have been measured extremely precisely down both boreholes, are the relics of former temperature fluctuations at the surface which have propagated into the ice sheet by diffusive heat flow and by ice flow. Working independently, two teams^{6,7} have assumed that the surface temperature history is described by the isotope profile but that the calibration is not known, nor necessarily linear. In an iterative procedure, they have then used an adjustable calibration factor to generate a surface temperature history which forces time-dependent, linked heat and ice-flow models until a best fit is achieved. An independent inverse solution for one of the boreholes has also been calculated (Dahl-Jensen).

The three approaches have independently resulted in very close matches with the measured borehole profiles, confirm-

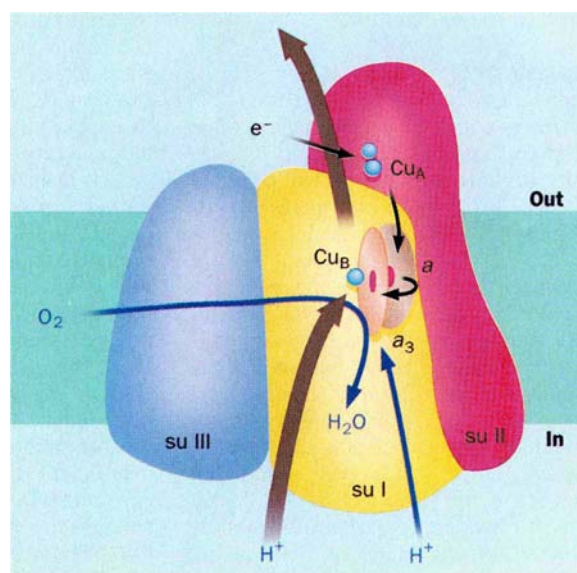
ing $\delta^{18}\text{O}$ as a reliable proxy for profiling at least the long-term average temperatures at Summit. They do indicate, however, that the calibration is time-dependent, with a considerably smaller gradient of isotope change with temperature in glacial times than today. A striking conclusion is that glacial to interglacial temperature changes were much larger than has previously been concluded from either ice cores or models. Climate in this region may be more sensitive than previously thought, reinforcing the evidence that changes in global temperatures may be amplified in the polar regions. Researchers engaged in predicting future sea level or modelling climate feedbacks arising from changes in polar albedo should take note. □

David A. Peel is at the British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK.

1. Mayewski, P. A. & Bender, M. *Rev. Geophys. Suppl.* 1287–1296 (July 1995).
2. GRIP project members *Nature* **364**, 203–207 (1993).
3. Grootes, P. M. et al. *Nature* **366**, 552–554 (1993).
4. Alley, R. B. et al. *Nature* **373**, 393–394 (1995).
5. Charles, C. D. et al. *Science* **263**, 508–511 (1994).
6. Cuffey, K. M. et al. *Science* (in the press).
7. Johnsen, S. et al. *Tellus* (in the press).

Correction

IN R. J. P. Williams's News and Views article of 24 August¹, which discussed the structure of cytochrome oxidase and associated details of proton and electron transfer, the incorrect impression was given that the accompanying diagram was based on the crystal structure data of Michel and co-workers² (the main paper under discussion). The diagram was in fact based on earlier work^{3,4} (see also refs 5,6). The figure shown here, kindly provided by Professor Michel, shows the appropriate picture from X-ray crystallography that will be the basis for future analysis of proton-electron coupling in cytochrome oxidase. The positions of prosthetic groups are different but much of the beta sheet and helical components are similar; as yet the indicated proton paths, especially for pumping, remain possibilities. The three subunits are labelled su I–III, and the haem groups *a* and *a*₃. In and out refer to the



mitochondrial matrix and intramembranous space respectively. Electron paths are as discussed in detail in ref. 2.

1. Williams R. J. P. *Nature* **376**, 643 (1995).
2. Iwata, S., Ostermeier, C., Ludwig, B. & Michel, H. *Nature* **376**, 660–669 (1995).
3. Wikström, M. *Nature* **338**, 775–778 (1989).
4. Williams, R. J. P. *Biochim. biophys. Acta* **1058**, 71–74 (1991).
5. Garcia-Horsman, J. A., Puustinen, A., Gennis, R. B. & Wikström, M. *Biochemistry* **34**, 4428–4433 (1995).
6. Williams, R. J. P. *Nature News and Views* **338**, 709–710 (1989).

A microbial minimalist

Barry R. Bloom

WHAT'S in a genome? Lots of genes, of course, but the first complete bacterial genome sequences to be published¹⁻³ reveal that each genome has a unique story to tell and adds its own new mysteries to be explained. The announcement in *Science* last month by Fraser *et al.*¹ of the *Mycoplasma genitalium* sequence comes on the heels of the *Haemophilus influenzae* (strain Rd) sequence released earlier this year^{2,3} by scientists from university and government and from the Institute of Genomic Research.

*Mycoplasma genitalium*¹ was chosen because it has the smallest genome known for a self-replicating organism (580 kilobases), with the hope that it might provide insight into the minimal functional gene set for a living organism. *Haemophilus influenzae* has a larger but still relatively small genome (1.8 megabases), and serves well as a model bacterium; the major human pathogen of this species is the type b, a major cause of otitis media, for which a most effective vaccine has recently become available. Although these are by no means the first complete genomes sequenced, the first being that of the phage ΦX174 (5,386 base pairs) by Sanger in 1977 and the most complex being that of cytomegalovirus (229 kilobases), they are the largest to date.

Sequencing

As physical maps did not exist for either bacterial genome, they were sequenced by a random shotgun strategy applied to the whole genome using 3–9-fold coverage and high throughput, and analysed with sophisticated assembly and alignment software^{2,3}. This means that, for *H. influenzae*, 28,000 sequencing reactions were performed; for *M. genitalium*, the entire genome was sequenced on eight sequencing machines by five individuals working for two months. Gaps were then closed by sequencing overlapping fragments obtained from independent libraries. This assault was remarkably rapid, error-free (1 base in 5,000–10,000), and cost-effective. A comparable amount of time was spent analysing the sequence data and sieving existing bacterial sequence information for comparison: likely genes were assigned and the genetic map annotated with putative operons, regulatory regions, starts and stops. The sequencing cost on this scale works out at 30 cents per base, or about \$200,000 for the complete sequence of the *Mycoplasma* genome.

What has been learned about the genes? In the case of *H. influenzae*, 1,743 possible coding regions were identified, 1,007 of which could take on various

functions in the cell, including amino-acid and lipid metabolism, biosynthesis of cofactors and the cell envelope, energy production, transport, nucleotide and protein synthesis, and DNA replication and transcription. The distribution of the respective open reading frames in *H. influenzae* is nicely shown in colour-coded maps, which give a handy picture of its genome organization and an indication of the amount of DNA a typical bacterium chooses to invest in different functions — for example, 10% goes to energy metabolism, 17% to transcription and translation, 12% to transport and 8% to cell envelope proteins.

In the case of the 470 predicted coding regions of *M. genitalium*, an organism that lives in association with mammalian cells, the investment is quite different. For example, whereas *H. influenzae* has 68 genes for amino-acid biosynthesis, *M. genitalium* has only one. It has no genes for cytochromes or enzymes of the tricarboxylic acid cycle. *Mycoplasma genitalium* is by no means the earliest eubacterium and it has obviously learned to simplify its life by association with its mammalian hosts, so it is fascinating to learn what it could afford to shed and still survive. Yet it has committed almost 5% of its genome to repeated elements encoding an adhesin, which presumably allows it to stick to the cells that nurture it and which can probably be altered by recombination to enable antigenic variants to elude the immune responses of its host.

This leads to some of the mysteries, the untold stories. Foremost is the fact that despite the wealth of information in the sequence databases from microorganisms, fully a third of the open reading frames of both genomes predict sequences that cannot be assigned any biological function. Does this mean that each organism has evolved some very specialized proteins, not common to others, to carry out important or unique functions, or simply that insufficient information is available to account for the species diversity? And what do the 90 genes found in *M. genitalium*, but not in *H. influenzae*, actually do? Finally, how does *M. genitalium* survive without a transcription factor for the stress response? Perhaps in simplifying its lifestyle it has learned how to avoid or cope with stress — perhaps a lesson for us all.

What does all this mean for microbiologists? At best, it means that they will be freed to do more biology, rather than just molecular biology. The need for random mutagenesis and screening — very inefficient approaches to defining gene functions even when transposable elements

are available, and overwhelming when not — will be superseded by amplification using the polymerase chain reaction of specific genes of interest. These will be of known or unknown function, readily mutated, and rapidly deleted from or inserted into the chromosome in order to approach the mysteries of uniqueness and diversity.

Availability

For pathogens, this means defining the genes that control virulence. It means that complete information on these — and scientists must insist that this applies to *all* sequenced genomes — should be available to scientists everywhere, not only in print, but in usable databases as in the work⁴ discussed here, to allow scientists to pursue the implications of this knowledge.

As with all microbial genetics, we can predict that some of the fall-out from knowing the complete genome sequences will enable organisms to be modified for medical, agricultural and economic purposes. And we must anticipate the more sinister application to biological warfare, calling for vigilant international surveillance.

Finally, in light of widespread public interest in the threat of infectious diseases and emerging pathogens, it is curious that in the initial formulation of the Human Genome Project not a single pathogen was included. Neither are the two organisms discussed here significant pathogens. Given the desperate need for public understanding and support of science, one could argue that, in the words of Tallyrand upon the assassination of the Duc de Broglie, "It was worse than a crime: it was a blunder".

The power and cost-effectiveness of modern genome sequencing technology mean that complete genome sequences of 25 of the major bacterial and parasitic pathogens could be available within five years. For about 100 million dollars (only 500 times the investment in the *Mycoplasma* genome), we could buy the sequence of every virulence determinant, every protein antigen and every drug target. It would represent for each pathogen a one-time investment from which the information derived would be available to all scientists for all time. And we could then think about a new, post-genomic era of microbe biology. □

Barry R. Bloom is at the Howard Hughes Medical Institute, Albert Einstein College of Medicine, Bronx, New York 10461, USA.

1. Fraser, C. M. *et al.* *Science* **270**, 397–403 (1995).

2. Fleischmann, R. D. *et al.* *Science* **269**, 496–512 (1995).

3. Smith, H. O., Tomb, J.-F., Dougherty, B. A., Fleischmann, R. D. & Venter, J. C. *Science* **269**, 538–540 (1995).

4. World Wide Web site <http://www.tigr.org>

Mutations caught in the act

Myron F. Goodman

SPONTANEOUS mutations, such as base substitutions and frameshifts, arise in DNA because DNA polymerases occasionally make mistakes. A central issue in mutagenesis is to determine the nature of mispaired intermediates that may occur in DNA replication. One possible answer lies in hypothetical structures involving disfavoured tautomers of the natural bases in Watson–Crick A:C and G:T mispairs¹. Tautomers are structures in which, as in the examples shown in the figure, a proton has been transferred to the opposite side of the usual hydrogen bond. On page 260 of this issue, Zewail and colleagues² have used the remarkable time-resolution of femto-second laser spectroscopy to follow the process of tautomerization in a model self-complementary base, 7-azaindole, in the absence of solvent.

The proposal that disfavoured tautomers of the natural bases (A_{imino} , C_{imino} , G_{enol} , T_{enol}) provide a structural basis for base substitution mutations was made in the classic paper on DNA structure by Watson and Crick¹. Until recently, it was the accepted explanation for transition mutations (substitution of purine by another purine and pyrimidine by another pyrimidine), and was contained in most of the popular texts on molecular genetics and biochemistry. But there has never been any convincing evidence demonstrating that disfavoured tautomers are responsible for either spontaneous or base-analogue-induced mutagenesis. To account for transversion mutations (substitution of purine by pyrimidine and vice versa), Topal and Fresco³ proposed a model in which one of the mispaired bases is a disfavoured tautomer and the other is in a *syn* conformation. However, the purine:purine mispairs observed by NMR with one of the bases in a *syn* conformation have structures retaining the favoured amino or keto tautomeric states⁴.

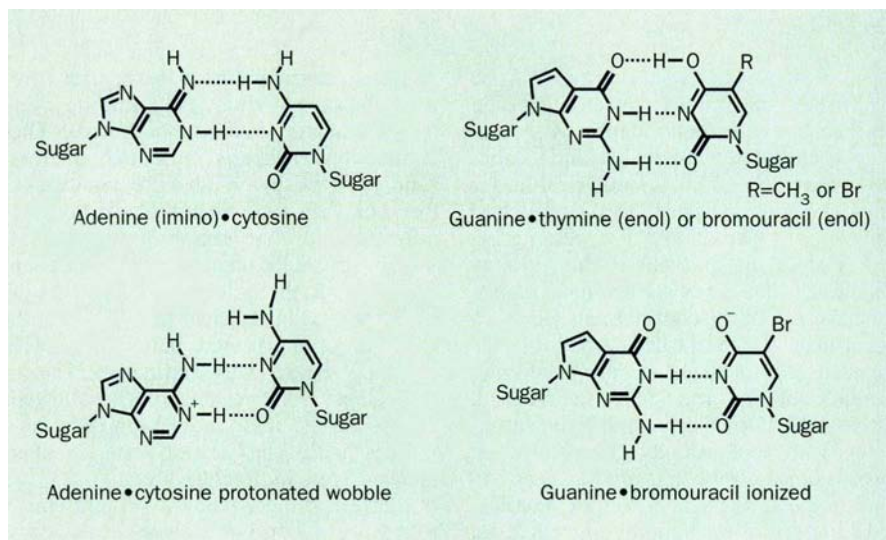
A model invoking ionization of base pairs was proposed by Lawley and Brookes⁵ to explain the mutagenic properties of 5-bromouracil (5-BU); in ionized form, 5-BU can form a two-H-bonded mispaired structure with guanine. Since 1986, NMR, X-ray crystallographic and enzyme kinetic studies have documented the presence of ionized and protonated base mispairs in DNA conforming to both Watson–Crick and wobble structures (the latter are where the position of paired bases is shifted in the aromatic plane relative to normal Watson–Crick base pairing, for example the structure found in G:U codon–anticodon recognition), with the mispaired

bases being present in their favoured amino and keto forms⁴.

The observation of ionized, protonated and wobble base-pair structures in DNA does not, of course, prove that such mispairs are formed in the active site of DNA polymerase. Perhaps the base pairs do exist in disfavoured tautomeric states in the polymerase active site and later shift to the structures observed by NMR and X-ray crystallography. Measurements of the incorporation kinetics for 5-BU:G mispairs and 5-BU:A correct

site constraints may discriminate against different base-mispaired structures.

The achievement made by Douhal, Kim and Zewail², then, is to have investigated a biologically important system in the gas phase by using femtosecond spectroscopy. For an isolated model base pair, the 7-azaindole dimer, they have measured timescales for formation of an intermediate excited state structure and its relaxation to form a tautomer, thus enabling the mechanism and dynamics of tautomerization to be established under near-vacuum gas phase conditions. Although the mechanism of ground state tautomerization is different from that induced by light, it is still of fundamental importance. Disfavoured tautomers have



Top, two examples of hypothetical structures proposed for A:C and G:T (or G:BU) base mispairs involving a disfavoured imino tautomer of adenine and enol tautomer of T (or BU). Bottom, the protonated wobble and ionized base-pair structures of the same mispairs observed by NMR studies of oligonucleotides in aqueous medium.

pairs as a function of pH have demonstrated catalysis of ionized base mispairs by DNA polymerase⁶. These data are inconsistent with a model in which disfavoured enol tautomers of either 5-BU or G are required to participate in the base mispairing event. Even so, the kinetic data cannot be used to rule out participation of disfavoured tautomers in mutagenesis.

The frequencies of incorporation of protonated, ionized or disfavoured tautomer base mispairs relative to correct base pairs depend on properties of the polymerase active site. Typical misinsertion frequencies fall in the range of 10^{-4} to 10^{-6} . Different DNA polymerases exhibit wide variations in nucleotide insertion fidelities⁴, perhaps because of differences in the degree of solvation of the active sites^{7,8}. The concentrations of ionized or disfavoured tautomeric species in the enzyme active site are likely to differ considerably from equilibrium concentrations in solution. Depending on the geometry of the base mispairs, enzyme active

been elusive structures, and the ability to identify them — interesting in itself — may open the way to investigate their properties in relation to ionized and wobble base pairs.

True, it is not the tautomer properties in vacuum but rather their properties in solution that must ultimately determine if they have any role to play in mutagenesis. Equilibria of base-mispaired structures and the relevant transition states in the polymerase active site are likely to be highly sensitive to solvation. Warshel and colleagues have proposed that enzyme

1. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 964–967 (1953).
2. Douhal, A., Kim, S. K. & Zewail, A. H. *Nature* **378**, 260–263 (1995).
3. Topal, M. D. & Fresco, J. R. *Nature* **263**, 289–293 (1976).
4. Echols, H. & Goodman, M. F. A. *Rev. Biochem.* **60**, 477–511 (1991).
5. Lawley, P. D. & Brookes, P. J. *molec. Biol.* **4**, 216–219 (1962).
6. Yu, H., Eritja, R., Bloom, L. B. & Goodman, M. F. *J. biol. Chem.* **268**, 15935–15943 (1993).
7. Petruska, J., Sowers, L. C. & Goodman, M. F. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1559–1562 (1986).
8. Warshel, A., Aqvist, J. & Creighton, S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5820–5824 (1989).

active sites are designed for electrostatic stabilization of ionic transition states by solvating these states to a greater extent than water can⁸.

At the time that Watson and Crick proposed their disfavoured tautomer model there were no known structures for base mispairs. Since then, NMR, X-ray and enzyme kinetic studies have revealed the presence of 'stable' mispairs consisting of ionized and wobble structures in favoured tautomeric states. So the ball is now in the other court — it is the mutagenic rele-

vance of disfavoured tautomers that is now in question. Are ionized and protonated base pairs sufficiently stable to 'short-circuit' the utilization of disfavoured tautomers in the polymerase active site? The study by Zewail and co-workers may open a window to address this issue. □

Myron F. Goodman is in the Department of Biological Sciences at the University of Southern California, University Park, Los Angeles, California 90089-1340, USA.

OPTOELECTRONICS

Silicon sees the light

David A. B. Miller

SILICON may sparkle in the eyes of the electronics community, but to those in optics it has remained a depressingly dull grey. It emits light very feebly, and its ability to control photons is similarly dim. Lu and colleagues, on page 258 of this issue¹, show that growing very thin, alternating layers of silicon and silicon dioxide may change all that. They see strong evidence that the quantum confinement that can occur in such 'superlattice' structures may allow us to turn silicon into a shining example of an optoelectronic material. Their superlattice can emit light (luminescence) with a colour that can be controlled by the choice of silicon layer thickness.

Optoelectronics is becoming increasingly important for handling information. Long-distance telecommunications are now dominated by optical fibres. Semiconductor lasers are key to compact discs for information storage. Lasers and light-emitting diodes are used in many printers, and in display devices such as liquid crystal panels and indicator lights.

With the notable exception of various kinds of silicon photodetectors, most semiconductor optoelectronic devices are made from compounds of group III and group V elements, such as gallium arsenide. These compounds dominate semiconductor optoelectronics because of the 'direct bandgap' of many III-V materials: their electronic structure is such that an electron can be boosted directly from the valence to the conduction band by absorption of a photon. Direct bandgaps result in strong optical absorption and efficient light emission; they permit the manufacture of compact high-speed photodetectors, bright light-emitting diodes, lasers and very effective optical modulators. Silicon, however, has an 'indirect bandgap' — a phonon (crystal vibration) is needed to complete the all-important optical transition involved in light absorption or emission, fatally weakening this optical transition for many applications.

When semiconductor structures are made in sizes of about 1–10 nm, the electrons become quantum-confined. The conduction electrons in semiconductors (and their positively charged analogues, the 'holes' in the valence band) are not only particles but quantum mechanical waves, so, when they are confined to a small region, the only allowed energies of electrons and holes are those corresponding to the standing wave patterns that will fit in this region. This confinement therefore allows some tuning of the allowed energy levels, and under some circumstances could change the bandgap of a material from indirect to direct.

Interest in the possibility of light emission from silicon got a boost a few years ago with the observation that a form known as 'porous silicon' luminesced brightly². The mechanism for this bright emission is still controversial³; although the structures are small, the emission wavelength does not show the dependence on size expected from quantum confinement. In contrast, the results obtained by Lu and colleagues¹ show size dependence very characteristic of quantum confinement.

Quantum confinement is already routinely used in III-V optoelectronics. Thin 'quantum well' layers give improved laser efficiencies and particularly effective optical modulators. Such quantum well or superlattice techniques have also been explored in silicon-germanium multilayered structures. Recently, for example, with additional quantum confinement from etching the layers into small posts (or 'quantum dots'), efficient, electrically driven light-emitting diodes have been demonstrated in silicon-germanium superlattices⁴. The approach of Lu and colleagues is different again, not only in its use of silicon dioxide, but also because the materials are not crystalline. The quality of their growth may be particularly crucial in making such amorphous materials luminesce efficiently. □

Why try to make a better silicon-based optoelectronic material? There may be some naivety in the community about the impact such a material would have, perhaps just as naive as the predictions made some years ago that III-V electronic devices might displace silicon electronics. III-V optoelectronics works well, can be low-cost and has a growing weight of technology behind it. Nonetheless, other options are always welcome, and the massive investment in the silicon industry could help a silicon-based technology. The choice by Lu and colleagues to work with silicon and silicon dioxide is particularly canny — this system must be one of the most investigated because of its central role in silicon electronics.

There is also the motivation of making optoelectronics work better with silicon electronics. As silicon electronics becomes ever more impressive (roughly doubling the number of transistors on a chip every 18 months, for example), the usual electrical interconnection technology of metal wires has a harder and harder job to send information within machines. Optical systems are already being used to replace long electrical cables and can also avoid many of the problems that plague wired connections at shorter distances (such as high power dissipation and limited bandwidth). The ability to integrate high-performance optoelectronics with silicon may be crucial in taking full advantage of the features offered by optical links. Growing III-V devices on silicon has proved difficult, but hybrid integration techniques are emerging, with thousands of III-V optoelectronic components successfully bonded to silicon circuits⁵. Silicon-based devices grown directly on silicon electronics could prove an attractive option.

There is, of course, much work to be done, and researchers into III-V optoelectronics can still sleep at night for some time to come. Nevertheless, the work of Lu and colleagues is an important demonstration of the potentially useful properties of these silicon-based structures. Many stages remain before we can seriously start to assess device possibilities (including, for example, electrically pumped light emission), but their work will surely stimulate more activity in this field and may yet give silicon an even brighter future. □

David A. B. Miller is in the Advanced Photonics Research Department, AT&T Bell Laboratories, 101 Crawfords Corner Road, Holmdel, New Jersey 07733, USA.

1. Lu, Z. H., Lockwood, D. J. & Baribeau, J.-M. *Nature* **378**, 258–260 (1995).
2. Canham, L. T. *Appl. Phys. Lett.* **57**, 1046–1048 (1990).
3. Hamilton, B. *Semiconductor Sci. Technol.* **10**, 1187–1207 (1995).
4. Tang, Y. S. et al. *Electr. Lett.* **31**, 1385–1386 (1995).
5. Goossen, K. W. et al. *IEEE Photonics Technol. Lett.* **7**, 360–362 (1995).

Out of Africa and into Asia

Bernard Wood and Alan Turner

FEW doubt that Africa was the birthplace of the hominid lineage, but there is no equivalent consensus about when hominids first moved out of that continent. Despite the announcement of early dates for a juvenile *Homo erectus* from Indonesia¹, the circumstances surrounding the recovery of many of the fossil hominids from the island will always hinder attempts to date them. Thus the excavation of hominid remains, in combination with crudely fashioned artefacts in what are claimed to be earliest Pleistocene deposits at Longgupo Cave in central China (Huang and co-workers, page 275 of this issue²), is of major importance. Most notably, the remains lend support to the idea³ that representatives of the hominid lineage were established in mainland Asia as early as about 1.9 million years (Myr) ago.

Africa has been the focus for research into human evolutionary history for the past three decades, but it was not always thus. A century ago, space in the correspondence columns of *Nature* was regularly claimed to debate the significance of the finds Eugene Dubois had made, beginning in 1891, at Trinil in Indonesia. Although initially allocated to *Pithecanthropus erectus*, the species distinction of the Trinil hominid has survived but the genus has long since been sunk into *Homo*.

Two decades later, excavations were instigated by the Canadian anatomist Davidson Black in the cave deposits at Choukoutien, now called Zhoukoudian, and the first of the series of remains of what became known as 'Peking Man' was

discovered. Despite being allocated to a new genus and species, their affinities with the hominids from Trinil, and with similar material that was subsequently recovered at Sangiran, also in Indonesia, was evident, and the Chinese remains have also been subsumed within *H. erectus*. There have been sporadic attempts to demonstrate both that the hominid remains from the Indonesian sites are from more than one species^{4,5}, and that they include specimens that should be allocated to *Australopithecus*⁶ or *Paranthropus*⁷, and thus to an earlier, more primitive phase of hominid evolution. But none of these claims has survived close scrutiny⁸. Likewise, until recently there has been little compelling evidence to suggest that any of the Asian hominid sites were yielding hominids more than one million years old⁹.

The importance of the material from Longgupo Cave is twofold. Not only does it support an early date for the hominid occupation of Asia, but the morphological details of the admittedly fragmentary fossil evidence also mean that it may represent not *H. erectus* but a more primitive species akin to *H. ergaster*, thus far known only from Africa.

Of course, dating the material is crucial to the argument. Longgupo Cave has several lines of evidence, none of them contradictory. Palaeomagnetic stratigraphy shows a reversed polarity for most of the sediments, with the hominid fossils and lithic items associated with the lower of two normal events and therefore referred to the Olduvai magnetic event. The magnetic evidence is broadly sup-

ported by analysis of tooth enamel from the sediments, using electron spin resonance, which gives a minimum age of 0.75 ± 0.09 Myr based on an early uranium uptake model. It *could* be argued that the normal magnetic event associated with the material is therefore likely to be Jaramillo, but the associated mammalian fauna is really too archaic and points instead to the earlier Olduvai event. Of particular interest here is the presence of *Nestoritherium*, a genus of the family Chalicotheriidae, an extinct, bizarre, claw-hoofed member of the Perissodactyla, today represented by tapirs, rhinos and horses.

The lithic items identified as primitive stone tools do seem to be exotic, and they are notably larger than the rest of the sediments. They look as much like stone tools as anything of this age ever does, and they fall into the category of items in finer sediment deposits that, as Gamble⁹ has pointed out, tend to categorize genuine archaeological assemblages as opposed to naturally bashed stones. Moreover, the uneroded state of the bone in clay facies channels is consistent with primary deposition rather than intrusive burial. But we are unlikely to be dealing with a site of hominid occupation. The giant hyaena, *Pachycrocuta*, is a perfectly plausible agent of accumulation¹⁰ (it is less likely that the sabre-toothed *Homotherium* did much bone destroying).

The authors draw attention to the presence of *Gigantopithecus*, a large, gorilla-like and presumably herbivorous primate, in the same level as the hominid fossils, and stress that this is the third such co-occurrence at Asian localities over a time span of some 1 Myr. Such co-occurrences are always intriguing, but the evidence of hyaena activity reduces the likelihood that *Gigantopithecus* was prey to the more

Australopithecus goes west

As discussed in the main article, on page 275 of this issue Huang *et al.* report evidence that bears upon the expansion of the range of hominids beyond Africa. But how well do we understand their distribution within that continent? Early hominid sites in Africa are concentrated in two regions: East Africa, where they are centred on the Gregory Rift Valley, and southern Africa, where fossil evidence has come from caves located in the high veldt. Discoveries in Malawi¹ have helped to bridge the gap between the two areas², but were the early hominids as restricted in their range as the distribution of these sites suggests?

Apparently not, for on page 273 Brunet *et al.*³ announce the discovery of an australopithecine-like mandible,

some 3–3.5 Myr old, from Chad. The known range of that genus is thus extended westward by 2,000 km or so.

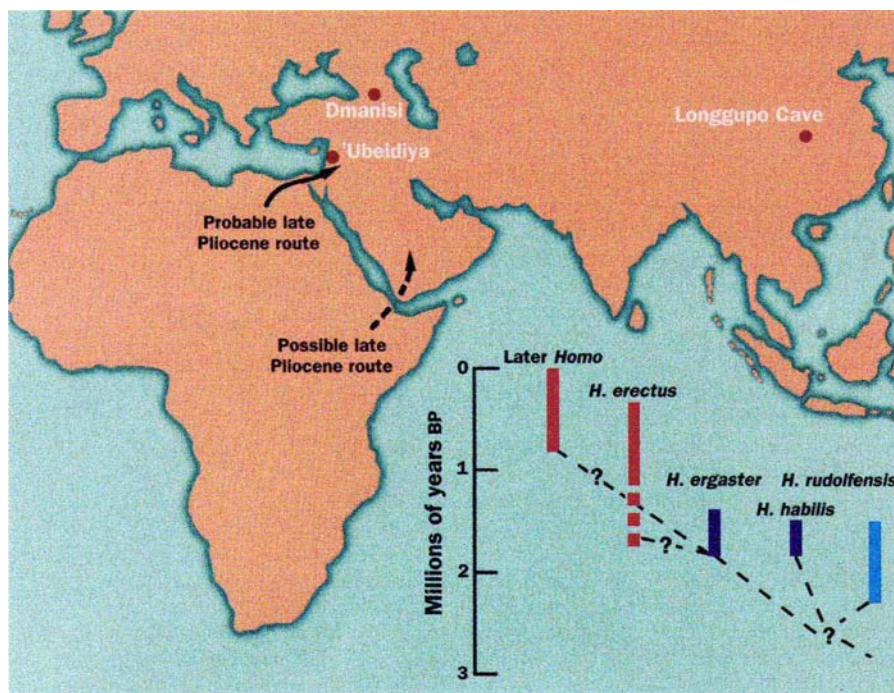
Determining the range of an extinct species is particularly tricky. What should the null hypothesis be? If the starting point is that the range is determined by the location of fossil discoveries, how does one accommodate the maxim that 'absence of evidence is not evidence of absence'? Even if suitable fossil sites exist, low population density and the vagaries of preservation will result in the representation of some species being very patchy. The Mio-Pliocene hyaenids are a good example⁴, and *Agriotherium*, a fossil bear, is only known from two sites in Africa⁵, but they are 6,000 km apart.

The hominid mandible from Chad is

clearly australopithecine in grade. It is closest in morphology to *Australopithecus afarensis*, but may turn out to be a new species; given its location, this would not be surprising. In the Pliocene, habitats similar to those indicated by the fauna associated with the Chad discovery extended from the Atlantic Ocean to the Cape. There is no reason to think that australopithecines did not use that range to the full.

B. W.

- Schrenk, F., Bromage, T. G., Betzler, C. G. & Ring, U. *Nature* **365**, 833–836 (1993).
- Wood, B. *Nature* **365**, 789–790 (1993).
- Brunet, M. *et al.* *Nature* **378**, 273–275 (1995).
- Werdelin, L., Turner, A. & Solounias, N. *Zool. J. Linn. Soc.* **111**, 197–217 (1994).
- Howell, F. C. in *Neogene Paleontology and Geology of Sahabi* (eds Boaz, N. T., El-Arnati, A., Gaziry, A. W., de Heinzelin, J. & Boaz, D. D.) 153–181 (Liss, New York, 1987).



The evidence from Longgupo Cave in China, described by Huang *et al.*² and discussed here, suggests that hominids were established in Asia just after two million years ago. Given the primitive nature of the premolar teeth, it seems that the first hominid to occupy Asia may not have been *Homo erectus*, but perhaps a variant of *H. ergaster* or even *H. habilis*.

advanced hominid. The remaining elements of the mammalian fauna at Longgupo shed little light on the local environment of the site, although both woodland and more open-country taxa seem to be represented.

The hominid remains — part of the left side of an adult mandible and an isolated upper incisor — are meagre pickings from a taxonomic point of view. However, the mandibular fragment includes both the crown and the root of a premolar tooth (P_4), and they provide the best evidence about the affinities of the material. The crowns of the P_4 teeth of *H. erectus* are generally relatively simple and the teeth are usually single-rooted, like those of modern humans. In contrast, the Longgupo P_4 root is bifid for most of its length. This, and other features of the mandible and the dentition, suggest that the Longgupo hominid may be much more primitive than *H. erectus*¹¹. This opens up the possibility that the first hominid to leave Africa was at least as primitive as *H. ergaster*¹², and implies that *H. erectus* may have evolved within Asia and

then spread back into Europe and Africa.

In terms of overall patterns of mammalian movement, there is nothing inherently implausible about the age of the material and the implications that it holds for human dispersions from Africa. Hominid remains and lithic items from Dmanisi¹³ in Georgia point to at least an initial presence at the gates of Europe by around the same time as the age of the Longgupo evidence. And it is clear that a Late Pliocene dispersion across Arabia, probably via the Levant, and perhaps through the Bab-el-Mandab straits, was possible for several mammalian taxa^{14,15}, while the presence of hominids in the Levant itself by 1.4 Myr ago is evident at the Israeli site of 'Ubeidiya¹⁴. The new report from Huang *et al.* adds weight to other, less well-substantiated claims that hominids travelled even further, and occupied China in the latest Pliocene some 1.9 Myr ago. □

Bernard Wood and Alan Turner are in the Department of Human Anatomy and Cell Biology, The University of Liverpool, PO Box 147, Liverpool L69 3BX, UK.

DAEDALUS

Flying boat

A MAGNETICALLY levitated monorail vehicle has a linear-motor system which generates a moving magnetic field beneath it. By moving endlessly downwards and backwards, the field repels a conducting track, levitating the vehicle and pushing it forwards at high speed.

Daedalus is now adapting the idea for nautical use. Salt water is a conductor, and could in principle support a magnetically levitated vehicle. It conducts so feebly, however, that a big, diffuse field would be needed, extending out and down through a huge volume of water. The small, tightly wound iron-cored or even superconducting magnets of normal maglev practice would be quite out of place here; huge, light, air-cored coils will be best for this job. The standard shore-based 50 Hz a.c. would be a poor choice of current too: sea water is a better conductor at high frequencies. Daedalus's system will work at many kilohertz. The upper limit will probably be set by distributed capacitance problems or radiation losses. DREADCO's engineers are already sketching designs for a wide, flat maglev hovercraft, only a few metres tall but with the area of an aircraft carrier, levitated and propelled by a mesh of wide, flat coils beneath its deck. With no wetted surface and no propellers, it will skim over the oceans silently, smoothly and very fast. In constricted waters, it will be amazingly manoeuvrable. Simple phase-shifts of its coils will send it forwards, backwards, sideways, or even spin it on its own axis.

The high speed and manoeuvrability of the new craft will be a boon to passengers. They may, however, find the voyage oddly disquieting. Largely salt water themselves, they will also experience the downward and backward drive of the craft's enveloping field, and may feel leaden, unbalanced, even glued to the decks — a useful safety feature in rough weather.

But rough weather should not be much of a problem. The maglev hovercraft must keep as close to the sea surface as possible, and Daedalus plans to achieve this by magnetically damping the waves. He is designing an interactive maglev system which, rather like dynamo braking on an electric train, captures the power of each rising wave-crest — thus flattening it out, and putting useful power into the coil above it. Fast sensing and switching between the coils will allow the outer ones to gain power by damping the waves, while the ones in the calm interior provide most of the levitation and thrust. Unlike all previous vessels, the maglev hovercraft will leave a calm wake behind it.

David Jones

1. Swisher, C.C. *et al.* *Science* **263**, 1118–1121 (1994).

2. Huang, W. *et al.* *Nature* **378**, 275–278 (1995).

3. Etler, D. A. & Li, T. in *Integrative Paths to the Past* (eds Corruccini, R.S. & Clochon, R. L.) 639–675 (Prentice Hall, New Jersey, 1994).

4. Jacob, T. in *The Origin of the Australians* (eds Kirk, R. L. & Thorn, A. G.) 81–93 (Humanities Press, New Jersey, 1976).

5. Sartono, S. in *Proc. 1st Congrès International de Paléontologie Humaine* (ed. de Lumley, H.) 491–533 (CNRS, Nice, 1982).

6. Von Koenigswald, G. H. R. *J. hum. Evol.* **2**, 487–491 (1973).

7. Robinson, J. T. *Am. J. phys. Anthropol.* **12**, 181–200 (1954).

8. Kramer, A. *J. hum. Evol.* **26**, 3–21 (1994).

9. Gamble, C. *Nature* **369**, 275–276 (1994).

10. Turner, A. & Anton, M. *Geobios* (in the press).

11. Wood, B. A. *et al.* *J. Anat.* **156**, 107–139 (1988).

12. Wood, B. A. *Nature* **355**, 783–790 (1992).

13. Gabunia, L. & Vekua, A. *Nature* **373**, 509–512 (1995).

14. Tchernov, E. in *Mammalian Migration and Dispersal Events in the European Quaternary* (eds Von Koenigswald, W. & Werdelin, L.) 103–123 (Cour. Forschungsinst. Senckenberg **153** (1992)).

15. Turner, A. *Rech. Archéol. l'Université de Liège* **62**, 9–17 (1995).

Sound attenuation by sculpture

SIR — It is generally accepted that sculptures or *objets d'art* are useless from the scientific point of view. We would like to show here an example of a certain type of sculpture that is of scientific relevance, and which can be related to the recent discovery of the so-called photonic band-gap materials and their generalization to other classical waves^{1,2}.

Theory predicts that a periodic distribution of wave scatterers, distributed in three-dimensional space, would induce severe attenuation of wave propagation in some spectral regions. The crucial parameters that allow the appearance of gaps are the ratio between the wave velocity in the scatterer and in the host, and also the volume fraction occupied by the scatterers³. For elastic or acoustic waves, the density ratio is also important. Experiments are mostly restricted to electromagnetic waves in the microwave region and gaps will appear in a region of wavelengths that corresponds to the periodicity of the scatterers and their geometric size. We are unaware of any experiments on band-gap materials for

acoustic waves where band gaps would appear in structures with a periodicity of between a few centimetres and one metre.

These types of structures are well known in modern art and are classified as a form of 'minimalism'. We report here on sound attenuation experiments performed on one of these sculptures (*a* in the figure). The sculpture, by Eusebio Sempere, is exhibited at the Juan March Foundation in Madrid. It consists of a periodic distribution of hollow stainless-steel cylinders, with a diameter of 2.9 cm, simple cubic symmetry and a unit cell of 10 cm. The cylinders are fixed on a circular, 4-m-diameter platform that can rotate around a vertical axis. We have measured sound attenuation in outdoor conditions for sound-wave vectors perpendicular to the cylinders' vertical axis.

The transmission characteristics in decibels vary as a function of the sound frequency (*b* in the figure) with the *k* vector along (100). Similar attenuation spectra have been obtained for other vectors perpendicular to (001), *k*_⊥. All the spectra have a large noise induced by the sound reflections from nearby buildings. Several maxima (sound attenuation) and minima (sound reinforcement) are present and their frequencies do not depend on *k*_⊥. The agreement between these features and theoretical maximum and minimum attenuation due to interference of the different crystal planes of the sculpture is fairly good, even though the sculpture does not have an ideal external shape for experimental purposes.

The Bragg attenuation peak at 1,670 Hz (which corresponds to [100] destructive interference) shows a ratio between the full width at half maximum and the peak energy position of 0.18. This value is similar to that obtained in photonic band experiments for a two-dimensional system⁴. The sculpture corresponds to a Cermat topology with a volume fraction occupied by the scatterers of 0.066 and a velocity ratio of 17.9. These values are close to those for which the calculations predict the appearance of a band gap for the propagation of sound waves in a two-dimensional periodic

structure³. Therefore, the sound attenuation peak at 1,670 Hz could be ascribed to the formation of the first gap in this sculpture.

**R. Martínez-Sala, J. Sancho
J. V. Sánchez, V. Gómez
J. Linares**

*Departamento de Física Aplicada,
Universidad Politécnica,
46022 Valencia, Spain*

F. Meseguer

*Instituto de Ciencia de Materiales (CSIC),
Campus de Cantoblanco,
28049 Madrid, Spain*

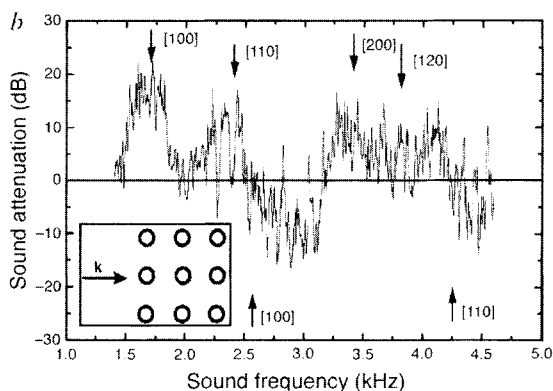
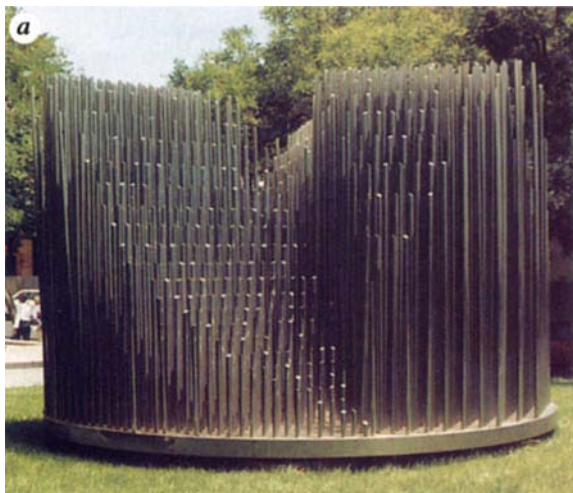
1. Yablouovitch, E. *Phys. Rev. Lett.* **58**, 2059–2062 (1987).
2. John, S. *Phys. Rev. Lett.* **58**, 2486–2489 (1987).
3. Economou, E. N. & Sigalas, M. M. *Phys. Rev.* **B48**, 13434–13438 (1993).
4. Smith, D. R. et al. *J. opt. Soc. Am.* **B10**, 314–321 (1993).

Why leaves are sometimes red

SIR — Anthocyanins are present in leaves of autumn foliage and rapidly developing shoots in tropical trees, but their function has never been satisfactorily explained. In addition, leaf undersurfaces of many tropical rainforest understory herbs are also red because of the presence of anthocyanins. We have examined leaves of two such taxa and report that anthocyanins intercept quanta otherwise absorbed by chlorophyll *b*, thereby protecting leaves from photoinhibition. We thus confirm what may be a general explanation for anthocyanin function in leaves.

The strong association of undersurface coloration with forest understory and the uniformity of anthocyanins suggest that they are selectively advantageous to plants in extreme shade. The hypotheses of increasing leaf temperature or protecting against damage by ultraviolet-B radiation have been discounted^{1,2}, and no mechanism has yet been discovered by which they could backscatter radiation and increase light-capture efficiency in the red wavelengths¹. We have thus examined the hypothesis of photoprotection by anthocyanins in these plants.

Anthocyanins *in vivo* should capture wavelengths otherwise absorbed by chlorophyll *b*, which has a longer wavelength absorbance in the Soret band. Chlorophyll *b* is especially associated with the light-harvesting complex of photosystem II, the principal site of photoinhibition and photo-damage in the chloroplast³. In understory shade, photosystem II-enriched plants are more susceptible to photoinhibition and damage in brief flecks of sunlight⁴. Anthocyanins should absorb at wavelengths otherwise absorbed by chlorophyll *b*, diminish photoinhibition and increase maximum photosynthesis. Anthocyanins may also allow higher chlorophyll concentrations, possibly correlated with changes in photosystem components and function.



a, Kinematic sculpture by Eusebio Sempere. *b*, Sound attenuation results as a function of the sound frequency. The wave vector is along the (100) direction as shown in the inset. Arrows indicate the calculated maxima and minima due to interference from the different crystal planes of the sculpture.

PIGMENT CONCENTRATIONS AND PHYSIOLOGICAL RESPONSES IN GREEN- AND RED-UNDERSURFACED LEAVES OF *BEGONIA PAVONINA* AND *TRIOLENA HIRSUTA*

Response	<i>B. pavonina</i>		<i>T. hirsuta</i>	
	Green	Red	Green	Red
Anthocyanin (n = 5)				
Absorption peak (nm)		530 ± 4		520 ± 9
µg cm ⁻²	0.0	14.0 ± 0.9	0.0	24.4 ± 4.1
% Dry mass	0.0	1.0	0.0	0.8
Chlorophyll (n = 10)				
a + b (µg cm ⁻²)	7.22 ± 0.22	10.05 ± 0.29***	17.37 ± 0.65	28.54 ± 1.22***
b (µg cm ⁻²)	1.82 ± 0.06	2.92 ± 0.10***	5.71 ± 0.19	9.94 ± 0.40***
a/b	2.96 ± 0.04	2.44 ± 0.04***	2.02 ± 0.02	1.86 ± 0.02***
% Dry mass		0.7		1.0
Maximum photosynthesis (µmol m ⁻² s ⁻¹) (n = 4)	1.87 ± 0.35	2.27 ± 0.34*	2.81 ± 0.25	3.80 ± 0.34*
F _v /F _m (n = 10)	0.672 ± 0.006	0.733 ± 0.007***	0.698 ± 0.011	0.780 ± 0.005***
Quantum efficiency				
Red-enriched (n = 4)	0.047 ± 0.014	0.038 ± 0.014 NS	0.026 ± 0.004	0.025 ± 0.002 NS
Far-red-enriched (n = 4)	—	—	0.030 ± 0.004	0.031 ± 0.002 NS

Means ± standard errors. *, **, *** Denote levels of significance of <0.05, 0.01 and 0.001, respectively; NS, not significantly different (Student's *t*-test); *n* = sample size. Plants were grown in a 12-h photoperiod at 25°C and at 35 µmol m⁻² s⁻¹ photosynthetic photon flux density (400–700 nm, PFD); *B. pavonina* in a red : far-red ratio of 0.12, and *T. hirsuta* in a ratio of 0.85. We isolated and identified anthocyanins², and estimated chlorophyll and anthocyanin concentrations^{10,11}. We measured transient fluorescence with a CF-1000 measurement system (P. K. Morgan Instruments, Andover, MA), with excitation at 500 mmol m⁻² s⁻¹ PFD and dark acclimation of 30 min. We measured photosynthesis with a Li-6200 photosynthesis system (Li-Cor Instruments, Lincoln, NE) illuminated by a tungsten-halogen lamp with heat-reflecting mirror and neutral-density filters, deriving apparent quantum efficiencies from the slopes of light-response curves with the same lamp, also filtered for far-red enrichment.

We examined red and green individuals from populations of two taxa native to rain-forest understory shade: *Begonia pavonina* Ridl. (Begoniaceae), from Bukit Lanjang Forest Reserve, Malaysia, and *Triolena hirsuta* Triana (Melastomaceae), from La Selva Research Station in Costa Rica. Undersurface reflectance⁵ above 570 nm was not greater in red than in green leaves of either species, which is inconsistent with the backscatter hypothesis¹. The anthocyanic mesophyll cells (with appreciable concentrations of peonidin-5-glucoside) overlapped in absorption wavelengths with chlorophyll *b* (see table)⁶. Anthocyanins in this layer should produce a spectral gradient that particularly protects the more shade-adapted chloroplasts deep within the leaf⁷. Variable to maximal fluorescence

(*F_v/F_m*; inversely correlated with photo-inhibition⁴) was significantly lower in green than in red leaves of both taxa. Anthocyanic leaves in *B. pavonina* sustained greater maximum photosynthesis, consistent with earlier measurements for *T. hirsuta*⁸. Higher chlorophyll concentrations explained the slightly greater red leaf absorbance at 700 nm (ref. 1). However, lower *a/b* ratios were not associated with an increase in the light-harvesting complex of photosystem II. Chloroplasts of red and green leaves of *T. hirsuta* had less photosystem-I and more photosystem-II oligomer than a spinach control, but were not different from each other from analysis by Deriphat polyacrylamide gel electrophoresis⁹. These leaves also had similar apparent quantum efficiencies (see table).

The photoprotective function of anthocyanins may pre-date their roles as visible pigments, as they occur in plant groups antecedent to the angiosperms. Thus, additional research may further our understanding of the evolutionary origin of anthocyanins.

Kevin S. Gould

*School of Biological Sciences,
University of Auckland,
Auckland, Private Bag 92019,
New Zealand*

David N. Kuhn

David W. Lee

Steven F. Oberbauer

*Department of Biological Sciences,
Florida International University,
and Fairchild Tropical Garden,
Miami, Florida 33199, USA*

Protecting HIV databases

SIR — HIV sequence data are accumulating at an ever-increasing pace. To maintain this information as a useful resource, it is imperative that viral sequence sets are carefully screened for problematic sequences before publication. Sources of error include sample mislabelling, viral culture contamination, or cross-contamination of reagents used for polymerase chain reaction (PCR) and molecular cloning¹. The most compelling indicators of potential problems include divergent viral sequences from a single individual or an epidemiologically linked pair, near-identity of epidemiologically unlinked viral sequences, or similarity to viral clones that are routinely used in the laboratory.

There are many examples of sequences that have been questioned, based on these indicators^{2–5}. For example, the original HIV-1 isolates were anomalously similar, and five publications were ultimately required to settle the issue of sample contamination⁶. Although such relationships may not necessarily be a consequence of experimental error, and instead reflect an interesting biological phenomenon, they nevertheless demand added scrutiny.

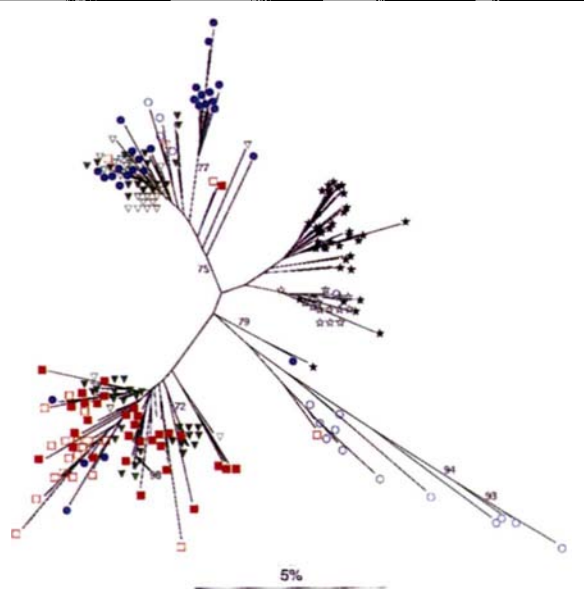
A case in point is the study by Briant *et al.*⁵, in which 308 HIV type-1 V3 envelope sequences were reported from four mother-to-infant transmission pairs. Each pair was evaluated separately, leading to the conclusion that viral sequences from three of the infants were more heterogeneous than their mothers' viral sequences, and that in these cases multiple viral lineages had been transmitted. These results contrast with the viral homogeneity typically observed early after infection in adults^{7,8} and in infants^{9–11}.

Close examination of the Briant *et al.* viral amino-acid sequence alignment (Fig. 4 in ref. 5) revealed instances of identity or near-identity in sequences from unlinked mother and infant pairs. Our subsequent examination of the viral nucleotide sequences aligned to a single consensus sequence (available on request) revealed several examples of such anomalies. Phylogenetic analysis of the combined sets of viral sequences revealed distinct clusters, yet sequences from each mother-infant pair were found in multiple clusters (see figure). Such evidence strongly suggests that contamination or mix-up of samples had occurred. Infant D was unique in showing little evidence of viral sequence contamination, and substantially lower within-patient viral diversity than the designated mother (see figure and ref. 5).

The quality of sequence data is of paramount importance if meaningful trends are to be found in the many studies undertaken

- Lee, D. W. in *On the Economy of Plant Form and Function* (ed. Givnish, T. J.) 105–131 (Cambridge Univ. Press, UK, 1986).
- Lee, D. W., Brammeier, S. & Smith, A. P. *Biotropica* **19**, 40–49 (1987).
- Anderson, J. M., Chow, W. S. & Oquist, G. in *Photosynthetic Responses to the Environment* (eds Yamamoto, H. Y. & Smith, C. M.) 14–26 (Am. Soc. Pl. Physiologists, Rockville, MD, 1993).
- Björkman, O. & Demmig, B. *Planta* **170**, 489–904 (1987).
- Lee, D. W. *Nature* **349**, 260–262 (1991).
- Lee, D. W., Bone, R. A., Tarsis, S. L. & Storch, D. *Am. J. Bot.* **77**, 370–380 (1990).
- Terashima, I. & Inoue, Y. *Pl. Cell Physiol.*, Tokyo **26**, 781–785 (1985).
- Fetcher, N., Oberbauer, S. F. & Chazdon, R. L. in *La Selva, Ecology and Natural History of a Neotropical Rain Forest* (eds McDade, L. A., Bawa, K. S., Hespeler, H. A. & Hartshorne, G. S.) 128–141 (Univ. Chicago Press, 1994).
- Peter, G. F., Takeuchi, T. & Thornber, J. P. *Methods: A Companion to Methods in Enzymology* Vol. 3, 115–124 (1991).
- Moran, R. *Pl. Physiol.* **69**, 1376–1381 (1982).
- Fuleki, T. & Francis, F. J. *J. Fd Sci.* **33**, 72–77 (1968).

Neighbour-joining phylogenetic tree for the viral sequences from the study of Briant *et al.*⁵. Filled and open symbols represent mother and child, respectively, for pair A (squares), pair B (triangles), pair C (circles) and pair D (stars). The scale bar represents 5% nucleotide sequence divergence. Phylogenetic relationships were estimated, using PHYLIP (J. Felsenstein, 1993; distributed on request by the author) for the 185 longest sequences (319–325 nucleotides) from the set of 308 sequences, using neighbour-joining from a matrix sequence distance using the two-parameter method of Kimura¹². Numerals at nodes indicate levels of bootstrap support of >70% of 400 bootstrap samples. The cases of sequence identity were not included in the analysis, and thus the bootstrap support of between-patient associations may be underestimated. A detailed description of this analysis will be submitted elsewhere.



to examine HIV evolution and the role of viral selection in transmission. Furthermore, these data, once submitted to sequence databases, are used to integrate data from multiple studies for further analysis. Therefore, a series of simple tests should be an integral part of viral sequence analysis protocols for identifying potential problems. New viral sequences should be compared with sequences of propagated viruses and molecular clones present in the laboratory where experiments are being conducted. Contamination would be revealed as identity or near-identity in a matrix of sequence similarities, or in the branch order and length observed in phylogenetic analysis. Neighbour-joining trees (as in the figure) are particularly well suited as a screening tool. They can reveal problematic associations swiftly and effectively even for large data sets. New viral sequences should also be screened against the sequences in GenBank.

It is the responsibility of investigators to ensure that their data are valid, and to discuss any potentially confounding results (as in ref. 4). A means of achieving this would be for journals to request that authors describe the methods they have

used to check for possible problems, and to provide the raw sequence data on request to facilitate the review process.

B. T. M. Korber*

*Theoretical Biology and Biophysics,
Los Alamos National Laboratory,
Los Alamos, New Mexico 87545, USA*

G. Learn

*J. I. Mullins
Departments of Microbiology and Medicine,
University of Washington,
K451 Health Sciences Center,
Seattle, Washington 98195-7740, USA*

B. H. Hahn

*Departments of Medicine and Microbiology,
University of Alabama at Birmingham,
701 S. 19th St,
Lyon-Harrison Research Bldg 613,
Birmingham, Alabama 35294, USA*

S. Wolinsky

*Department of Medicine,
Division of Infectious Diseases,
Northwestern University Medical School,
303 E. Chicago, Tarry Bldg 7-375,
Chicago, Illinois 60611, USA*

BRIANT *ET AL.* REPLY — We fully agree with Korber *et al.* concerning the importance of the issue of sample contamin-

ation. We had, in fact, carried out a more detailed analysis than theirs which was indeed considered as part of the review process of our paper⁵. Sample timing and order of processing demonstrate that no simple pattern of contamination, either of samples or of clones, can explain the observed sequence similarities.

A high level of similarity exists among some of the sequences from mother-child pairs A, B and C, although pair D sequences are quite distinct. The preparation of the peripheral blood mononuclear cell lysate for DNA purification was carried out at delivery as part of the diagnostic procedures. The dates of birth (A, 11/04/91; B, 07/12/91; C, 20/02/91; and D, 30/04/91) suggest that any problem with samples at this stage would affect pairs A and D, but these are clearly distinct. Second, the molecular analysis of the material (amplification, cloning and sequencing) was carried out in the order A, B, D, C, with the work on pairs B and D conducted simultaneously.

From the phylogenetic analyses we concluded⁵: for pair A no particular pattern was statistically supported; pair B had two distinct variants found in both mother and child — 'multiple transmission'; pair C's first child sequences were associated with early maternal sequences — early non-selective transmission; pair D's child sequences were associated with rare, late maternal sequences — selective perinatal transmission.

We have considered the impact of reasonable contamination scenarios on these conclusions. Pair B provides the strongest indication of multiple transmission because of the presence in the child of two distinct groups shortly after birth (5 weeks). Each child group was associated with a distinct maternal lineage. Although some pair B sequences were similar to sequences from pairs A and C, the key sequences from child B are quite distinct from those found in pair D, the sample analysed closest in time. They were also obtained sporadically among a background of the more common child form, which was not found in the sample processed earlier. There is no simple explanation for the multiple transmission on the grounds of cross-patient contamination. The evidence for early transmission in pair C is similarly robust — the relevant child sequences form a very distinct group close to HIV subtype-B reference isolates also analysed. Finally, the case of selective transmission involves the pair whose sequences are most distinct.

Sequence similarity is well established in HIV research. Clustered infections have been shown in several cases to have given rise to similar or identical sequences in different patients infected by both sexual and parenteral routes (refs 13–19 and D. Yirell, P. Robertson and A.J.L.B.,

1. Kwok, S. & Higuchi, R. *Nature* **339**, 237–238; Correction: **339**, 490 (1989).
2. McClure, M. O. *et al.*; Reply: Bryson, Y. & Chen, I. S. Y. *Nature* **375**, 637–638 (1995).
3. LaRosa, G. J. *et al.* *Science* **249**, 932–935 (1990); Clarifications: **251**, 811 (1991); **253**, 1146 (1991).
4. Antonoli, I. M., Baumberger, C., Yerly, S. & Perrin, L. *AIDS* **9**, 11–17 (1995).
5. Briant, L., Wade, C. M., Puel, J., Leigh Brown, A. J. & Guyader, M. *J. Virol.* **69**, 3778–3788 (1995).
6. Chang, S.-Y., Bowman, B. H., Weiss, J.B., Garcia, R. E. & White, T. J. *Nature* **363**, 466–469 (1993).
7. Zhang, L. Q. *et al.* *J. Virol.* **67**, 3345–3356 (1993).
8. Zhu, T. *et al.* *Science* **261**, 1179–1181 (1993).
9. Wolinsky, S. *et al.* *Science* **255**, 1134–1137 (1992).
10. Ahmad, N., Baroudy, B. M., Baker, R. C. & Chappey, C. J. *J. Virol.* **69**, 1001–1012 (1995).
11. Mulder-Kampinga, G. A. *et al.* *J. Virol.* **69**, 2285–2296 (1995).

12. Kimura, M. *J. molec. Evol.* **16**, 111–120 (1980).
13. Balfe, P., Simmonds, P., Ludlam, C. A., Bishop, J. O. & Leigh Brown, A. J. *J. Virol.* **64**, 6221–6233 (1990).
14. Simmonds, P., Balfe, P., Ludlam, C. A., Bishop, J. O. & Leigh Brown, A. J. *J. Virol.* **64**, 5840–5850 (1990).
15. Holmes, E. C. *et al.* *J. infect. Dis.* **171**, 45–53 (1995).
16. Cichutek, K. *et al.* *AIDS* **5**, 1185–1187 (1991).
17. Grez, M. *et al.* *J. Virol.* **68**, 2161–2168 (1994).
18. McCutchan, F. E. *et al.* *AIDS Res. hum. Retroviruses* **8**, 1887–1895 (1992).
19. Kalish, M. L. *et al.* *AIDS* **9**, 851–857 (1995).
20. Wain-Hobson, S. & Myers, G. *Nature* **347**, 18 (1990).
21. McNearney, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **87**, 1917–1921 (1990).

*Also at: The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 8750, USA.

manuscript in preparation). In a previous similar discussion, Wain-Hobson and Myers²⁰ cited the study of McNearney *et al.*²¹ as another example of sequence contamination. In fact, that paper was the first demonstration of the lack of variation observed in *env* at seroconversion, which has since become accepted as dogma in the field.

L. Briant

Laboratoire d'Immunologie des
Infections Rétrovirales,
CNRS UPR 9008 – INSERM U249,
Institut de Biologie,
34060 Montpellier, France

J. Puel

Laboratoire de Virologie,
Centre Hôpitalo-Universitaire Purpan,
31059 Toulouse, France

C. M. Wade

A. J. Leigh Brown

Centre for HIV Research,
Institute of Cell, Animal and Population

Biology,

Division of Biological Sciences,
The University of Edinburgh,
Edinburgh EH9 3JN, UK

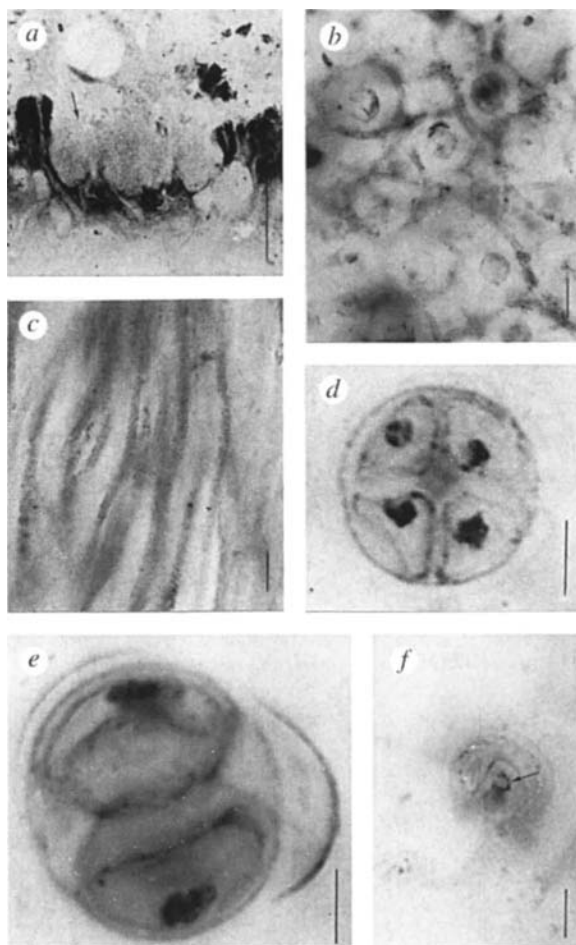
M. Guyader

Centre d'Immunologie,
Parc Scientifique et Technologique
de Luminy, Case 906,
13288 Marseille cedex 09, France

The oldest fossil lichen

SIR — We describe here the oldest unequivocal fossil lichen, from thin sections of Early Devonian (400 million years old) Rhynie chert. Lichens are symbiotic associations between a fungus (mycobiont) and a green alga or cyanobacterium (photobiont) which range from mutualism to controlled parasitism¹. The fungus gains a carbon source while the photobiont obtains protection, nutrients and water. Today, lichens can be found in a wide range of habitats, extending from close to the Poles, to deserts and rainforests.

The fossil thallus consists of thin bands approximately 1–2 mm thick, in which fungal hyphae are tightly aggregated (*a* in the figure). On the surface are numerous pockets containing hyphae in a three-dimensional net (*b* in the figure). Nets vary from circular to hexagonal, and average about 25 µm in diameter (*c* in the figure). Hyphae are aseptate, possess an irregular surface and are 1–4 µm in diam-



a, Longitudinal section of lichen thallus. Black areas are aggregated hyphae; the pocket (arrow) contains nets and cyanobacterial cells. *b*, Section of net showing hyphae surrounding photobiont cells. *c*, Detail of fungal hypha. *d*, Four of eight daughter cells within a common sheath. Note the degraded protoplast in each cell. *e*, Two daughter cells within a common sheath; outer sheath remnants are only partially preserved. *f*, Cyanobacterial cell showing possible site (arrow) where photobiont and mycobiont interact physiologically. Scale bars: *a*, 500 µm; *b*, 10 µm; *c*, 10 µm; *d*, 5 µm; *e*, 5 µm; and *f*, 5 µm.

eter. In the centre of each net is a solitary cell (15 µm diameter), or cluster of daughter cells (*d* in the figure), representing the photobiont. Each cell is surrounded by a thick envelope of mucilage. Some cells are surrounded by concentric sheaths, the remnants of the envelope (*e* in the figure), and these sheaths may contain up to 32 daughter cells. Many photobiont cells contain amorphous contents (*d–f* in the figure) which represent stages in protoplast shrinkage and degradation.

There is a consistent spatial relationship between cells of the photobiont and those of the mycobiont. At the base of each pocket the photobiont cells are small and solitary; cell size and the number of daughter cells increase towards the upper surface of the thallus. Hyphae are closely aggregated along the edge of the pocket and numerous filaments surround each photobiont cell. Towards the centre of the pocket, hyphae are less aggregated, and in the central region they are often dissociat-

ed from the cyanobacterial cells.

Some cells of the photobiont possess slight invaginations in the wall that may represent wall appositions or intragelatinous protrusions (*f* in the figure). These structures, and the fact that hyphae surround each of the photobiont cells, confirm the physiological interrelationship that existed between the two bionts. The fossil cyanobiont shares numerous morphological features with several extant cyanobacteria (for example, *Gloeocapsa*, *Chroococcidiopsis*), in which cells are enclosed in mucilage and produced in plates and spheres². All the daughter cell division patterns in the fossil can be found in extant gloeocapsid colonies. The mycobiont is more difficult to relate to modern fungi because no reproductive structures have been found.

Complex microbial communities were present around 3,500 million years ago³, but it is not known whether any consisted of physiological symbioses. To date, the oldest documented fungal symbioses are also found in the Rhynie chert in the form of endomycorrhizae⁴. The physiological interaction between the Devonian cyanobacterium and fungus resulted in the ability to exploit new ecological niches by competing

for space above ground, facilitating the exchange of gas and securing adequate light for the photobiont⁵. As a result, the Devonian cyanolichens were able to colonize and weather rock, thus contributing to the formation of soil, a necessary step in accommodating increasingly complex and larger terrestrial plants.

T. N. Taylor

Department of Botany,
University of Kansas,
Lawrence, Kansas 66045, USA

H. Hass, W. Remy

H. Kerp

Geologisch-Paläontologische Institut
und Museum,
Westfälische Wilhelms-Universität Münster,
4400 Münster, Germany

1. Ahmadjian, V. *The Lichen Symbiosis* (Wiley, New York, 1993).
2. Büdel, B. & Rhiel, E. *Archs Microbiol.* **143**, 117–121 (1985).
3. Awramik, S. M. *Photosynth. Res.* **33**, 75–89 (1992).
4. Taylor, T. N., Remy, W., Hass, H. & Kerp, H. *Mycologia* **87**, 560–573 (1995).
5. Honegger, R. *New Phytol.* **125**, 659–677 (1993).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words and five references.

The enigma of the natural killer cell

Jenny E. Gumperz & Peter Parham

Natural killer (NK) cells are controlled by receptors specific for polymorphic determinants of class I molecules of the major histocompatibility complex (MHC). The contrasting properties of NK and cytotoxic T cell (CTL) class I receptors provide complementarity in the cytolytic lymphocyte response to viruses, tumours and transplants. Whereas human NK cell class I receptors consist of immunoglobulin domains, their mouse counterparts resemble C-type lectins. This difference may reflect the receptors' diverse and rapidly evolving class I ligands.

MAMMALS use cytolytic T cells (CTL) and natural killer (NK) cells to eliminate tissues which are foreign, infected or in other ways compromised¹⁻⁷. Similarities in development, morphology, lytic mechanism and cell-surface phenotype indicate that NK cells and CTLs are closely related lineages⁸⁻¹¹. Their effector functions are controlled by MHC class I glycoproteins, and both cell types express receptors that bind these molecules. Whereas recognition of class I activates T-cell cytotoxicity, however, it usually turns off NK cells¹². CTLs and NK cells thus have complementary functions in the cytolytic response. Whereas CTLs recognize perturbed cells by the unfamiliar peptides their class I molecules present, NK cells look for the absence of class I, a common consequence of virus infection or malignant transformation.

Antigen recognition by CTLs is well understood. T-cell antigen receptors (TCRs), encoded by somatically rearranged genes¹³, engage ligands formed from short peptides bound to class I MHC molecules^{3,7,13}. MHC class I polymorphism (see Box) modulates the specificity of the peptide binding site and thus the foreign peptides which induce a CTL response¹⁴ (Fig. 1a). Our grasp of the basis for NK cell target choice is less firm¹⁵. The quest for a receptor analogous to the TCR that identifies targets for destruction has been fraught with complication. Numerous cell-surface glycoproteins contribute to the interactions between NK cells and their targets, but no single species stands out as 'the antigen receptor'. It has been more rewarding to examine the question from the target's stance: what makes a cell susceptible to NK attack? Evidence for the idea¹² that MHC class I molecules protect cells from NK cell lysis has been building, and it is now widely accepted¹⁶ (Fig. 2a, b).

Understanding this phenomenon has required comparison of target cells differing only in their class I allotypes, separating the inhibitory effect of class I from other variables. Cells from mice lacking class I expression owing to disruption of the β_2 -microglobulin gene, for example, are more susceptible to NK cell lysis than cells from normal mice^{17,18}, and the defect can be reversed by transfection with class I (refs 19, 20). Nevertheless, analysing the effect of class I molecules on NK cells remains difficult because the effect is negative—no lysis—and is controlled by MHC class I polymorphisms.

Polymorphic MHC class I determinants

The influence of MHC class I polymorphism was first appreciated from an observation that NK cell populations inhibited by other HLA-A and B allotypes were unperturbed by HLA-A*0201 (ref. 20). Changing residue 74 of the α_1 helix from histidine to aspartic acid, however, made HLA-A*0201 an inhibitor²¹. Subsequent studies using NK cell clones uncovered additional patterns of HLA class I-mediated inhibition^{22,23}. Substitutions at positions 77 and 80 of the α_1 helix divide HLA-C allotypes into two groups inhibiting separate subsets of NK cell clones^{24,25}. In an overlapping region (residues 77–83), certain HLA-B allotypes have a sequence motif correlated with Bw4, a serological epitope²⁶, defining a third inhibitory specificity²⁷

(Fig. 1b), and a substitution at position 80 within the Bw4 motif defines a fourth related specificity²⁸; additional specificities may also exist^{23,29}.

Polymorphic class I determinants also inhibit mouse NK cells. Mouse NK cells expressing the cell-surface marker Ly-49A are inhibited by targets expressing the class I molecules H-2^d and H-2^k, for example, unlike non-expressors³⁰. Rat NK cells, however, are stimulated by class I epitopes. Five stimulatory specificities map to polymorphisms of RT1-C, a class I gene distinct from RT1.A which presents antigens to CTL^{31,32}.

The polymorphic determinants that inhibit human NK cells are shared by subsets of HLA-B or C allotypes, and lie in the carboxy-terminal half of the α_1 helix (Fig. 1b). Although less precisely defined, the inhibitory determinants of mouse NK cells include groups of class I allotypes and map to the α_1 and α_2 domains^{30,33}. By contrast, CTLs recognize substitutions throughout the α_1 and α_2 domains, usually defining a single allotype^{13,14} (Fig. 1a). Their indifference to most class I polymorphism and their focus on the α_1 helix suggest that NK cell receptors do not confront the 'top' of the class I molecule as the TCR is thought to do¹³, but interact instead with the 'side', as do immunoglobulin Fc and the class I-like Fc receptor (FcRn)³⁴, or bacterial superantigens and class II MHC molecules³⁵.

Mouse NK class I receptors are lectin-like

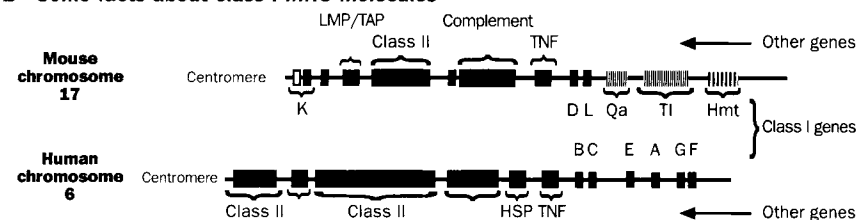
As suggested by its expression on NK cells inhibited by H-2^d and H-2^k (ref. 30), Ly-49A is a receptor for H-2^d and H-2^k class I allotypes^{36,37}. At the time of this observation, much was already known of the structure and genetics of Ly-49 genes³⁸ which form part of a larger multigene family on mouse chromosome 6 (refs 39–41) (Table 1). Ly-49C is now known to be a class I receptor specific for H-2^{b,d,k+s} (refs 42, 43). Ly-49 genes encode membrane proteins with extracellular domains resembling calcium-dependent (C-type) lectins^{44–46}. Ly-49A indeed binds sulphated mono- and polysaccharides, and the former reduce NK cell inhibition by H-2D^d. It is unclear whether binding of Ly-49A to H-2D^d requires glycosylation of the class I molecule⁴⁷ but asparagine 86, the conserved glycosylation site of MHC class I heavy chains, lies next to the residues defining human NK cell specificities (74–83).

Another gene family closely linked to Ly-49 also encodes C-type lectins involved in NK cell function. The region containing the two gene families has been called the NK gene complex (NKC)⁴⁸ and is responsible for NK cell-mediated resistance to cytomegalovirus infection⁴⁹. The founding member of the second family is NKR-P1, a molecule first identified on rat NK cells that transduces a stimulatory signal^{50,51}, exhibits calcium-dependent binding of sugars⁵², and stimulates lysis of otherwise resistant tumour-cell targets when crosslinked⁵³. The mouse NKC encodes three NKR-P1 homologues, one of which encodes the NK1.1 antigen which marks all NK cells of certain mouse strains^{54,55}. It also encodes CD69, another lectin-like molecule which marks early NK and T-cell activation (Table 1)⁵⁶.

Human chromosome 12 is the homologue of mouse chromosome 6 and encodes proteins homologous to the lectin-like

BOX 1 Some facts about class I MHC molecules

DIAGRAM showing simplified MHC organization in humans and mice. MHC class I molecules are heterodimeric peptide binding proteins, consisting of a variable transmembrane heavy chain encoded by genes in the MHC together with β_2 -microglobulin, which is usually non-polymorphic and encoded by a single copy gene not linked to the MHC. The number of heavy chain genes in the MHC varies from six in domestic pigs⁸⁶ to more than a thousand in the African pygmy mouse *Nannomys setulosus*⁸⁷ (including both functional genes and pseudo-genes). Between one and three heavy-chain genes are expressed codominantly on most tissues and cell types, and they exhibit extensive allelic polymorphism in populations. Humans and mice have three such genes: HLA-A, B and C in humans, and H-2D, K and L in mice. The number of HLA-A, B, C allotypes so far discovered in human populations exceeds 200, and high-resolution DNA-based methods for HLA typing are discovering new alleles every week⁸⁸. Most modern humans are heterozygous and therefore express 2 HLA-A, 2 HLA-B and 2 HLA-C allotypes. Class I heavy-chain alleles differ by 1–77 nucleotide substitutions in a total coding sequence of 1,100 nucleotides, and the substitutions are biased towards alterations in the peptide binding site or on the



Line diagram showing in simplified form the gene organization of mouse and human. Whereas orthologies can be discerned between class II MHC genes and other MHC genes in the two species, no such orthologies can be distinguished for the class I genes. All mouse class I genes are more closely related to each other than to any human class I gene, 200, and high-resolution DNA-based ¹⁰¹and vice versa.

surfaces of the α helices contacting the TCR^{89,90}. Within the α_1 and α_2 domains forming the peptide binding site, allotypes differ by an average of 16 substitutions⁸⁸. Each allotype binds many different peptides derived by intracellular degradation of proteins made within the cell, but the residues permitted at 1–3 'anchor' positions in the sequence are constrained by polymorphic specificity pockets within the class I peptide binding site⁷. Each class I allotype is thought to be able to bind thousands of peptide species. Class I MHC polymorphism is generally believed to stem from positive darwinian selection to diversify the range of peptides displayed by an individual⁸⁹.

molecules of the murine NKC⁵⁷. These include NKR-P1 and CD69 homologues^{58,59} and the gene for CD94 (ref. 60), another NK cell antigen. Although these genes apparently define a human NKC, none correspond to the mouse Ly-49 subfamily. Instead the human NKC contains a further family of four lectin-like genes (NKG2; refs 57, 61) with no mouse counterparts⁴⁵.

Although subfamilies of NKC-encoded lectin-like proteins exist, they differ from the C-type lectins encoded elsewhere in the genome^{45,46,62}. The NKC-encoded genes thus appear to have evolved by progressive gene expression and divergence from a common ancestor, giving rise to different subfamilies, not all of which occur in every species or strain of the same species. Similar patterns occur in MHC class I subfamilies (see Box).

Human NK class I receptors

Monoclonal antibodies that bind to subsets of human NK cells and interfere with their class I-mediated inhibition have been instrumental in identifying the class I receptors of human NK cells^{63–66}. By binding to a receptor, such antibodies prevent class I binding, allowing the NK cell to kill its prospective target as though the class I molecule were not there (Fig. 2c). The polymorphic specificities defined in this way correspond to those defined by class I HLA inhibition of NK cell clones^{25,27,67}.

The antibodies recognize glycoproteins of relative molecular mass 58,000 or 70,000 (M_r 58K or 70K) (Table 2), which represent a previously undiscovered family within the immunoglobulin superfamily comprising at least four genes on chromosome 19 (refs 68–70). No evidence for somatic rearrangement or mutation of the type seen in CTL receptors¹³ has emerged. The encoded proteins have an extracellular portion of either 2 or 3 immunoglobulin constant-like domains, a transmembrane anchor, and a cytoplasmic domain of varying length

(Fig. 3). Receptors that bind HLA-C have two immunoglobulin-like domains⁶⁸, whereas those recognizing the Bw4 epitope have three^{69,70}. Each functionally defined class I specificity is represented by more than one cDNA sequence, suggesting that further functional diversity will be correlated with molecular heterogeneity, as has already been suggested for Bw4 specificity^{27,28}.

Each human NK cell clone expresses at least one receptor, and some express two^{68,69}. Typing NK cell populations with receptor-specific antibodies also reveals cells expressing two distinct receptors⁷¹, which can be inhibited by class I allotypes recognized by either receptor. The observed patterns of inhibition are consistent with certain NK cells expressing three class I receptors. The implications for the control of receptor gene expression may only emerge once it is clear whether different genes, allelic polymorphism, alternate mRNA splicing, or other mechanisms are responsible for receptor variation.

Divergent mouse and human NK receptors

NK cells share both the capacity to kill a variety of target cells and a lack of memory with the cytolytic cells of lower vertebrates and invertebrates, suggesting that they are innate and primitive components of the vertebrate immune system^{6,8,12,72}. If NK cells are 'living fossils' of ancient origin, the molecular mechanisms defining their identity should predate the divergence of modern mammalian species.

The differences between the class I receptors of human and mouse NK cells (Fig. 3) challenge this paradigm. Have primates and rodents acquired different systems for class I-mediated regulation of NK cells? Such notions of convergent evolution are the choice of last resort, as they challenge the antiquity and generality of any inhibitory mechanism. More popular is the view that lectin- and immunoglobulin-like class I receptors are present

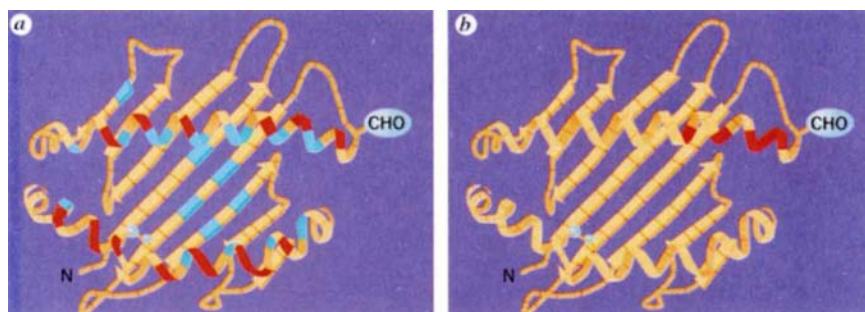


FIG. 1 Ribbon diagram of the α_1 and α_2 domains of the MHC class I molecules showing the positions of polymorphism that affect CTL interaction (a) and NK cell interaction (b). In a, residues that directly contact TCR are shown in red, and those that indirectly affect TCR contact through altering the spectrum of peptides bound are shown in blue.

TABLE 1 Genes localized to the NK gene complex (NKC)

	Mouse: chromosome 6			Human: chromosome 12			
	Ly-49	NKR-P1	CD69	NKR-P1	CD69	CD94	NKG2
Names of antigens	Ly-49A, B, C, D, E, F, G, H	NKR-P1A, B, C or NK1.1	VEA	hNKR-P1A	AIM or EA-1 or Leu 23	KP43	?
Multigene family	Yes	Yes	No	No?	No	No	Yes
Structure	S-S dimer	S-S dimer	S-S dimer	S-S dimer	S-S dimer	S-S dimer	?
M_r of one protein subunit	~30–33K	~26K	~24K	~25K	~23K	~20K	~25K
M_r of glycosylated dimer	~85–110K	~80K	~64K	~80K	~60K	~70K	?
Lectin activity	Yes	Yes	?	?	?	?	?
Ligands	Sulphated and phosphorylated CHO, class I molecules	Glycosaminoglycans, gangliosides, and sulphated CHO	?	?	?	Class I molecules?	?
Functions	Inhibition of target cell lysis	Stimulation of target cell lysis	Cellular activation, proliferation	Regulation of lysis	Cellular activation, proliferation	Stimulation/inhibition of lysis?	?
Expressed by	NK, T cells	NK, T cells	Haemopoietic cells	NK, T cells	Haemopoietic cells	NK, T cells	NK, T cells

Genes localized to the NK complex (NKC). All code for proteins which possess an extracellular domain with homology to C-type lectins, and which are predicted to have type II orientation in the membrane. The human and murine CD69 and NKR-P1 proteins share 58% and 47% amino acid identity, respectively. Members of the NKG2 family are most closely related to CD69 and NKR-P1. CD94 does not have significantly homology to any of the human or murine genes, except members of the NKG2 family (27–33%) (refs 16, 23, 35, 41–46, 48–66).

in both species, but that the mouse immunoglobulin-like and human lectin-like receptors await discovery^{68,73}.

In pondering this puzzle it is worth remembering the receptors' ligands. The view is striking: class I MHC molecules of humans and mice differ so much that no shared ancestor between individual genes or alleles in the two species can be discerned^{74,75} (see Box). The α_1 helix epitopes recognized by the human receptors are absent in mice and even in some primates. For example, HLA-C, the locus which may contribute most to human NK cell regulation, appears to be confined to humans and apes⁷⁶. Given the differences in the ligands, the homologues of human class I receptors in mice and vice versa (if any) could be quite different.

Individual mice³⁰ and humans (J.E.G., unpublished observations) commonly express receptors for which they have no ligand, suggesting that populations and species may on occasion lose the ligand for a particular NK receptor. Subsequent deletion of the receptor gene would then represent a neutral change, perhaps accounting for the presence of individual NK cell receptor genes (or even families) in one mammalian species that are absent in another.

Class I-mediated regulation of NK cells

In humans and mice, class I molecules generally turn off NK cells by delivering a negative signal^{77–79}. Human NK cell receptors are also expressed by subpopulations of T cells, and when bound to class I similarly inhibit T-cell function^{80,81}. Class I can on occasion stimulate NK cell function^{31,82}, and for rats this seems the

TABLE 2 NK cell class I receptors

	Ly-49	p58	NKB1
Species	Mouse, rat	Human	Human
Multigene family	Yes	Yes	Yes
Localization to NKC	Yes	No	No
Structure	S-S dimer	Monomer	Monomer
Protein type	Type II integral membrane protein	Type I integral membrane protein	Type I integral membrane protein
Protein superfamily	C-type lectin	Immunoglobulin	Immunoglobulin
Protein M_r (glycosylated)	30–33K (44–54K)	38K (55–58K)	50K (70K)
Ligands	H-2 molecules Carbohydrate	HLA-C molecules	HLA-B molecules

NK cell class I receptors. Murine and human class I receptors are multigene families which are similar functionally, but have little structural commonality. The Ly-49 gene family maps to murine chromosome 6, site of the NK gene complex. Genes encoding the human receptors map to chromosome 19 (refs 35, 41–43, 48, 52, 53, 67–70, 72, 73).

general rule, adding to the intrigue of species differences³¹. Such qualitative differences in the NK cell response to class I ligands parallel T-cell biology, where TCR engagement can trigger either positive or negative selection in the thymus, and activation or anergy in the periphery¹, depending on circumstances.

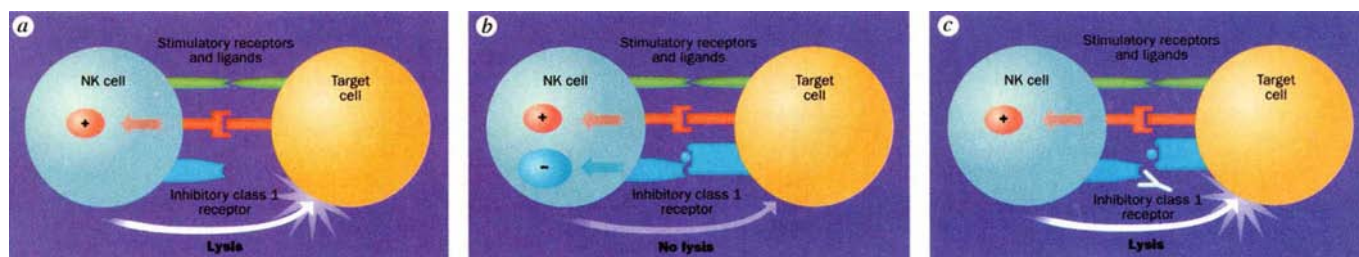


FIG. 2 Schematic representation of the interaction between NK cells and potential targets. a, A target cell that does not express MHC class I molecules is shown being killed. The signals generated by engagement of stimulatory receptors lead to the generation of an overall positive signal (+) that turns on the lytic machinery of the NK cells. b, An otherwise similar but class I-expressing target cell is not lysed by the NK cell,

owing to engagement of a class I receptor. This engagement leads to generation of a dominant negative signal (–) that prevents activation of the lytic machinery. c, Addition of a monoclonal antibody with specificity for the NK receptor interferes with engagement of class I MHC molecules and generation of the negative signal. Here the target cell is lysed.

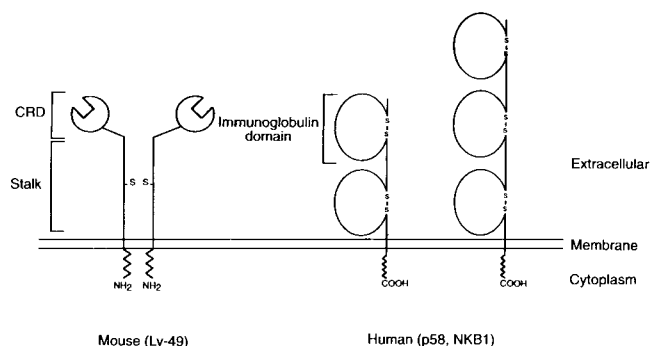


FIG. 3 Schematic representation of the differences in primary structure between the lectin-like class I receptors of the mouse and the immunoglobulin-like receptors of the human.

It has been proposed that NK cells complement CTLs by taking charge in situations where class I expression is compromised through the effects of virus infection or malignant transformation^{12,83}. NK cells would thus attack cells with reduced class I expression, sparing healthy cells that express normal levels of class I. By using several class I receptors, each specific for a different polymorphic class I epitope and expressed by different subsets of NK cells, at least one subset will be sensitive to the loss of any individual allotype and perceptive of 'missing self'⁸⁴. This strategy also decreases the likelihood of a pathogen mimicking all recognized epitopes and inhibiting all available NK cells.

If NK cells attack upon noticing loss of 'self' class I molecules, then an individual's NK cells must express at least one receptor

inhibited by one of the class I allotypes expressed by that individual. The NK cell repertoire must thus be driven or selected by self class I. By analogy with T cells, maturation of developing NK cells in the bone marrow may depend on engagement of class I receptors. Mice deficient or transgenic for MHC class I molecules offer some evidence supporting this model, as the phenotype and function of their NK cell populations is perturbed^{18,85}. From an analysis of Ly-49A, the effect appears to be quantitative, in that the number of NK cells expressing a receptor is unaffected, but the level of receptor expression reflects the levels of the MHC class I ligand⁸⁵.

Concluding remarks

Commonly perceived as nonspecific in function, primitive in nature, and defined by what they do not have and cannot do, NK cells have been overshadowed by B and T cells. Has their turn for the limelight arrived? NK cells, like CTLs, respond specifically to polymorphic determinants of MHC class I molecules, and the two cell types have many properties suggesting a complementary partnership in the cytolytic immune response.

The known class I receptors of mouse and human NK cells are so different that mechanisms for regulating their function must have evolved independently. Whether this evolution took place before or after separation of the two species is debated. The former hypothesis predicts that both receptors will be found in humans and mice, the latter that receptor families will show the lineage specificity as seen for class I MHC genes. Whatever the situation, the changes in NK cell regulation during mammalian evolution are hardly the adaptations of a 'living fossil'. □

J. E. Gumperz and P. Parham are in the Departments of Structural Biology and Microbiology & Immunology, Sherman Fairchild Building, Stanford University, Stanford, California 94305, USA

1. Sprent, J. in *Fundamental Immunology* (ed. Paul, W. E.) 75–109 (Raven, New York, 1993).
2. Trinchieri, G. *Adv. Immun.* **47**, 187–386 (1989).
3. Townsend, A. & Bodmer, H. A. *Rev. Immun.* **7**, 601–624 (1989).
4. Welsh, R. M. *Nat. Immun. Cell Growth Reg.* **5**, 169–199 (1986).
5. Yu, Y. Y., Kumar, V. & Bennett, M. A. *Rev. Immun.* **10**, 189–213 (1992).
6. Herberman, R. B. & Ortaldo, J. R. *Science* **214**, 24–30 (1981).
7. Barber, L. & Parham, P. A. *Rev. Cell Biol.* **9**, 163–206 (1993).
8. Herberman, R. B., Reynolds, C. W. & Ortaldo, J. R. A. *Rev. Immun.* **4**, 651–680 (1986).
9. Sanchez, M. J., Spits, H., Lanier, L. L. & Phillips, J. H. *J. exp. Med.* **178**, 1857–1866 (1993).
10. Sanchez, M. J. et al. *J. exp. Med.* **180**, 569–576 (1994).
11. Lanier, L. L. et al. *J. Immun.* **137**, 2735–2739 (1986).
12. Kärre, K., Ljunggren, H. G., Piötek, G. & Kiessling, R. *Nature* **319**, 675–678 (1986).
13. Davis, M. M. & Björkman, P. J. *Nature* **334**, 395–401 (1988).
14. Björkman, P. J. & Parham, P. A. *Rev. Biochem.* **59**, 253–288 (1990).
15. Herberman, R. B. & Ortaldo, J. R. A. *Rev. Immun.* **2**, 359–394 (1984).
16. Trinchieri, G. *J. exp. Med.* **180**, 417–421 (1994).
17. Liao, N. S., Bix, M., Zijlstra, M., Jaenisch, R. & Raulet, D. *Science* **253**, 199–202 (1991).
18. Bix, M. et al. *Nature* **349**, 329–331 (1991).
19. Shimizu, Y. & DeMars, R. *Eur. J. Immun.* **19**, 447–451 (1989).
20. Storkus, W. J. et al. *J. Immun.* **143**, 3853–3857 (1989).
21. Storkus, W. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5989–5992 (1991).
22. Ciccone, E. et al. *J. exp. Med.* **172**, 47–52 (1990).
23. Litwin, V. et al. *J. exp. Med.* **178**, 1321–1336 (1993).
24. Colonna, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7983–7985 (1992).
25. Colonna, M. et al. *Science* **260**, 1121–1124 (1993).
26. Wan, A. M., Ennis, P., Parham, P. & Holmes, N. J. *Immun.* **137**, 3671–3674 (1986).
27. Gumperz, J. E. et al. *J. exp. Med.* **181**, 1133–1144 (1995).
28. Cella, M. et al. *J. exp. Med.* **180**, 1235–1242 (1994).
29. Ciccone, E. et al. *J. exp. Med.* **175**, 709–718 (1992).
30. Karhofer, F. M., Ribaud, R. K. & Yokoyama, W. M. *Nature* **358**, 66–70 (1992).
31. Vaage, J. T. et al. *J. exp. Med.* **180**, 641–651 (1994).
32. Rolstad, B., Wonigelt, K. & Vaage, J. T. in *Natural Immunity to Normal Hemopoietic Cells* (ed. Rolstad, B.) 99–149 (CRC, Boca Raton, FL, 1994).
33. Sentman, C. L. et al. *J. Immun.* **153**, 5482–5490 (1994).
34. Burmeister, W. P., Huber, A. H. & Björkman, P. J. *Nature* **372**, 379–383 (1994).
35. Jardtzyk, T. S. et al. *Nature* **368**, 711–718 (1994).
36. Daniels, B. F. et al. *J. exp. Med.* **180**, 687–692 (1994).
37. Kane, K. P. *J. exp. Med.* **179**, 1011–1015 (1994).
38. Yokoyama, W. M. & Seaman, W. E. A. *Rev. Immun.* **11**, 613–635 (1993).
39. Yokoyama, W., Kehn, P., Cohen, D. & Shevach, E. *J. Immun.* **145**, 2353–2358 (1990).
40. Wong, S. et al. *J. Immun.* **147**, 1417–1423 (1991).
41. Smith, H. R., Karhofer, F. M. & Yokoyama, W. M. *J. Immun.* **153**, 1068–1079 (1994).
42. Sentman, C. L., Hackett, J. Jr., Kumar, V. & Bennett, M. J. *exp. Med.* **170**, 191–202 (1989).
43. Brennan, J., Mager, D., Jeffries, W. & Takei, F. *J. exp. Med.* **180**, 2287–2295 (1994).
44. Yokoyama, W. et al. *J. Immun.* **143**, 1379–1386 (1989).
45. Chambers, W. H., Adamkiewicz, T. & Houchins, J. P. *Glycobiology* **3**, 9–14 (1993).
46. Drickamer, K. *Prog. nucleic Acid Res. molec. Biol.* **45**, 207–232 (1993).

47. Daniels, B. F. et al. *Immunity* **1**, 785–792 (1995).
48. Yokoyama, W. M. et al. *J. Immun.* **147**, 3229–3236 (1991).
49. Scalzo, A. A. et al. *J. Immun.* **149**, 581–589 (1992).
50. Giorda, R. et al. *Science* **249**, 1298–1300 (1990).
51. Ryan, J. C. et al. *J. Immun.* **147**, 3244–3250 (1991).
52. Bezouska, K. et al. *J. biol. Chem.* **269**, 16945–16952 (1994).
53. Bezouska, K. et al. *Nature* **372**, 150–157 (1994).
54. Giorda, R., Weisberg, E. P., Ip, T. K. & Trucco, M. J. *Immun.* **149**, 1957–1963 (1992).
55. Ryan, J. C. et al. *J. Immun.* **149**, 1631–1635 (1992).
56. Ziegler, S. F. et al. *J. Immun.* **152**, 1228–1236 (1994).
57. Yabe, T. et al. *Immunogenetics* **37**, 455–460 (1993).
58. Lanier, L. L., Chang, C. & Phillips, J. H. *J. Immun.* **153**, 2417–2428 (1994).
59. Lopez-Cabrera, M. et al. *J. exp. Med.* **178**, 537–547 (1993).
60. Chang, C. et al. *Eur. Immun.* (in the press).
61. Houchins, J. P., Yabe, T., McSherry, C. & Bach, F. H. *J. exp. Med.* **173**, 1017–1020 (1991).
62. Adamkiewicz, T. V. et al. *Immunogenetics* **39**, 218 (1994).
63. Moretta, A. et al. *J. exp. Med.* **172**, 1589–1598 (1990).
64. Moretta, A. et al. *J. exp. Med.* **171**, 695–714 (1990).
65. Litwin, V. et al. *J. exp. Med.* **180**, 537–543 (1994).
66. Melero, I. et al. *J. Immun.* **152**, 1662–1673 (1994).
67. Moretta, A. et al. *J. exp. Med.* **178**, 597–604 (1993).
68. Wagtmann, N. et al. *Immunity* **2**, 439–449 (1995).
69. Colonna, M. & Samaridis, J. *Science* **268**, 405–408 (1995).
70. D'Andrea, A. et al. *J. Immun.* **155**, 2306–2310 (1995).
71. Lanier, L. L. et al. *J. Immun.* **154**, 3320–3327 (1995).
72. Versteeg, R. *Immun. Today* **13**, 244–247 (1992).
73. Barinaga, M. *Science* **268**, 367–368 (1995).
74. Rogers, J. H. *EMBO J.* **4**, 749–753 (1985).
75. Parham, P. *Semin. Immun.* **6**, 373–382 (1994).
76. Watkins, D. I. in *Critical Reviews in Immunology* (ed. Atassi, Z.) (CRC) (in the press).
77. Correa, L., Corral, L. & Raulet, D. H. *Eur. J. Immun.* **24**, 1323–1331 (1994).
78. Kaufman, D. S., Schoon, R. A. & Leibson, P. J. *J. Immun.* **150**, 1429–1436 (1993).
79. Kaufman, D. S. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 6484–6488 (1995).
80. Phillips, J. H., Gumperz, J. E., Parham, P. & Lanier, L. L. *Science* **268**, 403–405 (1995).
81. Ferrini, S. et al. *Eur. J. Immun.* **24**, 2294–2298 (1994).
82. Yu, Y. Y. et al. *Int. Immun.* **6**, 1297–1306 (1994).
83. Ljunggren, H. G. & Kärre, K. *Immun. Today* **11**, 237–244 (1990).
84. Ohlen, C. et al. *Science* **246**, 666–668 (1989).
85. Olsson, M., Kärre, K. & Sentman, C. *Proc. natn. Acad. Sci. U.S.A.* **92**, 1649–1653 (1995).
86. Kindt, T. J. & Singer, D. S. *Immun. Res.* **6**, 57–66 (1987).
87. Delarbre, C., Jaulin, C., Kourilsky, P. & Gachelin, G. *Immunogenetics* **37**, 29–38 (1992).
88. Parham, P., Adams, E. J. & Arnett, K. L. *Immun. Rev.* **143**, 141–180 (1995).
89. Hughes, A. L. & Nei, M. *Nature* **335**, 167–170 (1988).
90. Björkman, P. J. et al. *Nature* **329**, 506–512 (1987).

ACKNOWLEDGEMENTS. We thank V. Ogbuehi for help in preparing this manuscript. This work was supported by grants from the NIH.

Magnetism of the carbon allotropes

R. C. Haddon

The difference in magnetic susceptibility of graphite and diamond prompted Raman to postulate the flow of currents around the ring system of graphite in response to an applied magnetic field. The discovery of new carbon allotropes, the fullerenes, has furthered our understanding of this phenomenon and its relationship to aromatic character. C₆₀ and the other fullerenes exhibit both diamagnetic and paramagnetic ring currents, which exert subtle effects on the magnetic properties of these molecules and provide evidence for the existence of π -electrons mobile in three dimensions.

THE concept of molecular ring currents was developed from a consideration of the experimental magnetic susceptibilities of the two allotropes of carbon known in 1929. Raman¹ noted that: "It is known the graphite possesses a high specific susceptibility, ... quite ten times larger than the specific susceptibility of diamond, the latter being practically the same as that of carbon in organic compounds as found from Pascal's additive law. The abnormal susceptibility of graphite becomes intelligible in terms of the peculiar structure of the substance and its electrical conductivity, if we assume that there are electron orbits circulating round the plane hexagonal rings of carbon in the crystal lattice. ... Diamond, on the other hand, being a dielectric would not show the abnormal susceptibility". This explanation, which remains as the basis for our understanding of the difference in magnetic properties of diamond and graphite (Fig. 1), was based on an earlier paper by Ehrenfest² in which he rationalized the very large diamagnetism of bismuth on the basis of electron orbits that include several atoms within their radius.

These papers together motivate the continuing interest in the magnetic properties of conjugated organic molecules and materials. Although it is usual to focus on the localized electrons of diamond and the delocalized electrons of graphite, the magnetic properties provide a length scale for the electronic circulations induced by an applied magnetic field and show that in certain compounds of carbon interatomic currents can flow in the same manner as they do in a metal like bismuth.

In an extension of the Ehrenfest model, Pauling³ and Lonsdale⁴ applied the Larmor–Langevin formula⁵ to the π -electrons of aromatic organic molecules. The magnetic moment due to the field is given by the product of the current and the area of the loop over which it flows. Where all electronic shells are filled, the (molar) magnetic susceptibility due to the Larmor precession of an electron in an atom is given by

$$\chi = -\frac{Ne^2}{4mc^2} \rho^2 \quad (1)$$

where $\pi\rho^2$ is the area of the current loop, ρ^2 is the mean square of the distance of the electronic circulation from the axis of the magnetic field (z), N is Avogadro's number, c is the speed of light, and m and e are the electron mass and charge, respectively. For a spherically symmetric distribution of charge of radius r (where $r^2 = x^2 + y^2 + z^2$)

$$\rho^2 = (2/3)r^2 \quad (2)$$

With this substitution, equation (1) describes the magnetic susceptibility of atoms, and when one inserts the radius of the electronic circulations these equations reproduce the experimental values of the magnetic susceptibilities of closed-shell atoms⁵. This term (equation (1)) is always diamagnetic because the induced circulation gives rise to a moment that opposes the applied field, thereby raising the energy of the atom (the atom is repelled by the magnetic field). The early workers^{1–4} adapted equation (1) to the π -electrons of aromatic organic molecules

by solving for the mean square radius of a two-dimensional electron flow that is confined to the x - y plane (as in a circular wire of radius r)

$$\rho^2 = x^2 + y^2 = r^2 \quad (3)$$

But this procedure represents a severe approximation in the non-spherical potential of a molecule, because the quantum-mechanical theory of molecular magnetism⁵ shows that in this case an additional term is introduced into the equation for the susceptibility. Specifically, in addition to a diamagnetic term analogous to equation (1), there is a paramagnetic (so-called Van Vleck) term. This term arises because the nonspherical potential causes mixing of excited states into the electronic ground state by magnetic dipole transitions; this mixing has the effect of lowering the energy of the molecule and is therefore paramagnetic (the molecule is attracted by the magnetic field). In the spherical potential of an atom, the Van Vleck term vanishes, and the quantum-mechanical treatment reduces to the classical result embodied in equation (1).

Pauling³ began by considering the geometry of aromatic molecules and, working under the assumption that π -electrons were free to circulate around the six-membered rings (equations (1) and (3)) and thereby ignoring the Van Vleck term, he was able to rationalize the experimental susceptibilities of a number of these molecules. As discussed below, this 'classical' treatment provides a good estimate for the ring-current magnetic susceptibility of benzene, indicating that the assumption of a cylindrical potential for the π -electrons (equation (3)) works well in this case. From this perspective, the pronounced π -electron ring-current effect in benzene arises because the Van Vleck paramagnetic term in this system is small and can be neglected, rather than because the diamagnetic term is large. Lonsdale⁴, in an equivalent treatment, used the experimental susceptibilities to deduce the area and radius of the electronic circulations induced by the magnetic field (equations (1)–(3)). The argument by Lonsdale is appealing because she first calculates the atomic diameter of the atoms in aliphatic compounds and shows that they are compatible with the values given by X-ray diffraction, in spite of the approximations involved in such a treatment. She goes on to derive effective radii for the electronic circulations for carbon compounds of 0.7 Å (diamond), 0.8 Å (aliphatic), 0.7 Å (aromatic σ -electrons), 1.5–1.6 Å (aromatic π -electrons) and 7.8 Å (graphite π -electrons), an area of about 30 benzene nuclei. The first three values are clearly appropriate for electronic circulations that are confined to atoms, whereas the latter two must involve current flows that encompass a number of atoms and this requires that the currents flow along the bonds and around the conjugated rings; hence the term 'ring currents' is appropriate because the currents can only travel around closed cycles. The figure for the aromatic π -electrons is in good agreement with the radius of a benzene ring. This analysis clearly shows the rationale behind the continuing interest in the magnetism of conjugated carbon compounds: properly interpreted, the results

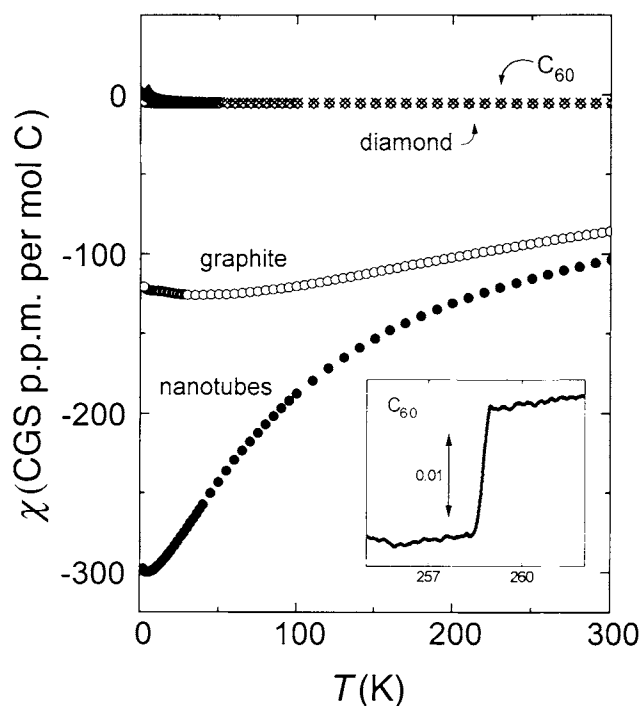


FIG. 1 Magnetic susceptibilities of the carbon allotropes as a function of temperature (CGS p.p.m. mol⁻¹ = e.m.u. × 10⁻⁶ mol⁻¹ = cm³ × 10⁻⁶ mol⁻¹). The inset shows the change in magnetic susceptibility in solid C₆₀ at the molecular orientational-ordering transition at 259 K. Adapted from Ramirez *et al.*⁴⁸.

provide an index of the length scale of the cycle of electronic circulations induced by an applied field. Notwithstanding Lonsdale's success in treating the magnetism of carbon compounds using equation (1), it is now standard practice to allow for additional, localized contributions to the susceptibility associated with atomic circulations, within a constitutive scheme of the type introduced by Pascal⁵. These contributions remain roughly constant in all carbon compounds (including diamond, graphite and the fullerenes). In the remainder of the article the empirical estimation of localized contributions is taken for granted, and I focus on the magnetism of the π -electrons.

Based on this discussion it is also clear why these electronic circulations rapidly became one of the panoply of characteristics expected of aromatic compounds—that is, one of the criteria that define aromatic character: they provide information on the ability of the electrons to move along the bonds in cyclic molecules. The properties expected of aromatic compounds (the nature of their chemical reactivity, their stability, structure and magnetism)^{6,7} have long superseded the original definition of aromaticity (smell). These attributes relate to the conjugated system as a whole and depend on the non-local character of the π -electrons—the ability of the π -electrons to act at a distance and to respond to the topology of a cyclic system. The resonance energy of such systems, and their tendency to undergo substitution reactions in which the character and symmetry of the system is retained, depend on the special stability associated with the presence of $(4n+2)\pi$ -electrons in a closed ring.

In this Review I will show what the third class of allotropes of carbon—the fullerenes—have added to our understanding of magnetically induced interatomic electronic circulations (ring currents). Like their two-dimensional counterparts, the fullerenes show evidence for ring currents but on the length scale of the individual rings, and the currents can be paramagnetic as well as diamagnetic (like those in benzene and graphite). The subtlety of fullerene magnetism has only recently become fully appreciated, and it is now recognized that the NMR spectra of

endohedral and exohedral derivatives, as well as of the doped forms, are a strong function of this property. Nevertheless the magnetic properties of the fullerenes are not a simple function of cluster size, as had been widely anticipated.

The magnetic properties of fullerenes

In 1985, Kroto *et al.* reported⁸ the observation of C₆₀ and the fullerenes, and their paper included the statement: "... the chemical shift in the NMR of the central atom should be remarkable because of the ring currents." Subsequently Smalley said⁹: "The structure results in a molecule that can almost be viewed as spherical benzene ... the molecule is aromatic, essentially a hexagonal graphite sheet, which, by incorporating pentagons curls up into a ball." Thus the question of the magnetism and what this might indicate for the aromatic character of the third allotrope of carbon was raised immediately upon its discovery.

In 1987 Elser and Haddon^{10,11} addressed this problem with a new solution of the London equations that considerably eased the treatment of topologically complex non-planar conjugated organic molecules. They also considered what magnetic properties would be expected of C₆₀ on the basis of the free-electron or Larmor-Langevin treatment of electronic circulations used by Pauling³ and Lonsdale⁴ (equations (1)–(3)). What would be the length scale of the electronic circulations induced by a magnetic field? If we define $\chi_{RC}^{\perp}$ to be the π -electron ring-current contribution to the molar susceptibility with field perpendicular to the molecular plane, the Pauling free-electron treatment gives³

$$\chi_{RC}^{\perp}(\text{benzene}) = -6 \frac{Ne^2}{4mc^2} l^2 \quad (4)$$

where l is the radius of the benzene ring, equal to the length of a carbon-carbon bond (about 1.4 Å). This formula gives³ $\chi_{RC}^{\perp}(\text{benzene}) = -49$ CGS p.p.m. mol⁻¹, in remarkable agreement with modern estimates of this quantity (even better agreement is attained with the exact area). If indeed C₆₀ were a spherical benzene with 60 π -electrons free to precess in a magnetic field, just as the electrons do in an atom, then the length scale for electronic circulations in equation (1) would be the radius of the sphere (r), in which case $\rho^2 = 2r^2/3 = 4.09l^2$ (equation (2)). Thus $\chi_{RC}(C_{60})/\chi_{RC}^{\perp}(\text{benzene}) = (60/6)(4.09)$. Hence C₆₀ would have 41 times the ring-current magnetic susceptibility of benzene, leading to a value of $\chi_{RC}(C_{60}) \approx -1,700$ CGS p.p.m. mol⁻¹—an enormous diamagnetism¹⁰. This would imply a theoretical value of $\chi(C_{60}) \approx -30$ CGS p.p.m. per mol C, to be compared with the measured magnetic susceptibilities of the other carbon allotropes of $\chi(\text{diamond}) = -5$ and $\chi(\text{graphite}) = -88$ in the same units (Fig. 1).

But the London calculations^{10,11} gave an entirely different result: $\chi_{RC}(C_{60}) \approx 0$, for equal bond strengths (Fig. 2). The vanishingly small ring-current contribution was ascribed to a large Van Vleck paramagnetism that cancels almost exactly the diamagnetic term; as indicated earlier, the former is neglected in equation (1). I indicated previously that equation (1) correctly describes the ring-current magnetism of benzene because the π -electrons move in a roughly cylindrical potential (equation (3)); the failure of this treatment for C₆₀ implies that the π -electron potential is far from spherical (as equation (2) assumes). Presumably the constraint that the currents must follow the network of bonds in the truncated icosahedral structure leads to the large Van Vleck term. Elser and Haddon¹⁰ found that the π -electron ring-current magnetic susceptibility was very sensitive to the relative bond strengths in the molecule and estimated that $\chi_{RC}(C_{60})/\chi_{RC}^{\perp}(\text{benzene}) \approx 0.3$ (Fig. 2) for realistic bond lengths and thus $\chi_{RC}(C_{60}) \approx 10$ to -120 CGS p.p.m. mol⁻¹.

Thus the third allotrope of carbon posed fundamental questions in molecular magnetism. Whereas many planar aromatic organic compounds could be successfully treated with either the classical free-electron theory or the quantum-mechanical London theory (which often gave similar results)¹², this was clearly not the case for the fullerenes, for which the classical and Lon-

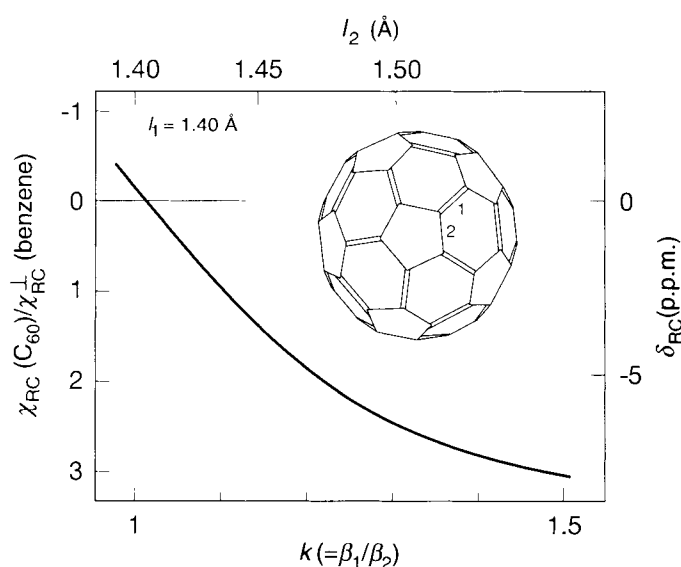


FIG. 2 Ring-current magnetic susceptibility (relative to benzene) and chemical shift for a central atom of C_{60} as a function of bond alternation and bond length, calculated with the London theory, where l_1 and β_1 refer to the bonds at the junction of two 6-MRs, and l_2 and β_2 apply to the bonds at the junction of 5- and 6-MRs (l and β are the bond lengths and resonance integrals, respectively). Adapted from Elser and Haddon^{10,11}.

don predictions differed significantly. If the London calculations were correct, did this imply that there were no interatomic electronic circulations and no ring currents in C_{60} ? And if so, how could this be rationalized in the light of the presumed aromatic character of these species and the very large ring-current diamagnetism of graphite?

Once the isolation and large-scale synthesis of fullerenes was achieved in 1990 by Krätschmer *et al.*¹³, measurements of the magnetic susceptibility were reported soon afterwards: $\chi(C_{60}) = -260$ CGS p.p.m. mol⁻¹ = -4.3 CGS p.p.m. per mol C (Fig. 1)^{14,15}. If we follow Lonsdale⁴, then the length scale for electronic circulations in C_{60} on the basis of these measurements (Fig. 1) (using equation (1)) is 0.6 Å—comparable to that found in materials in which it is known that the electronic currents are confined to the atoms. Another way of saying this is that the localized contributions in C_{60} alone account for the experimentally observed magnetic susceptibility. From this standpoint the electronic currents induced in C_{60} by a magnetic field must be viewed as even more localized than those in diamond, for which $\chi(\text{diamond}) = -5.5$ CGS p.p.m. per mol C. In fact, if C_{60} had been discovered in Pascal's day it would have represented a striking confirmation of his additivity law (that the molecular susceptibility is the sum of the localized atomic contributions), although difficult to rationalize with other known aromatics and graphite. These measurements provided a definitive verification of the London theory, and showed that the Van Vleck paramagnetic term is very important in C_{60} (refs 10, 14). This is as one might expect given that the HOMO → LUMO transition is magnetic-dipole-allowed¹⁵.

This finding left an important role for C_{70} to play, because an alternative explanation for the vanishingly small ring-current magnetic susceptibility in C_{60} would have been to propose that ring currents do not exist in the fullerene allotropes simply because, unlike graphite, they are not planar. The Hückel rule¹⁶ states that: "Planar, monocyclic conjugated systems are aromatic if they contain $(4n + 2)$ π -electrons", but an essential precursor to this famous rule was the extension by Hückel of the separability of σ - and π -electrons from one to two dimensions—from linear to planar conjugated systems¹⁷. Apparently orbital orthogonality provides a theoretical framework for σ - π separ-

ability in three dimensions, at least in some cases¹⁸. A basis set of π -orbitals may be defined in any dimension by constructing each atomic π -orbital so that it is orthogonal (of zero overlap integral) to the σ -orbitals at the conjugated carbon atom¹⁸. The observation of ring currents in the fullerenes would provide unequivocal evidence for the existence of π -electrons in three dimensions. Experimental measurements^{14,15} showed that the molar magnetic susceptibility of C_{70} is about twice that of C_{60} , indicating that diamagnetic ring currents must be present in C_{70} , and that their contribution to the magnetism is not fully cancelled by the Van Vleck term. This was the first evidence that ring currents play a role in fullerene magnetism and provided further support for the London calculations, which suggested¹⁴ that $\chi_{RC}(C_{70})/\chi_{RC}(\text{benzene}) \approx 6-7$.

Fullerene ring currents

From the preceding discussion of fullerene magnetism it is clear that the primary problem in the detection of ring currents in molecules by measurement of the bulk magnetic susceptibility is the need to first account fully for the localized atomic contributions that are present in all molecules^{19, 21}. In the case of C_{60} , the whole of the experimentally observed magnetic susceptibility could be accounted for by local atomic circulations, and even for benzene the local contribution accounts for 75% of the observed molar susceptibility. But it was shown by Pople²² that ring currents produce clearly discernible magnetic shielding effects on the NMR chemical shifts of protons in the vicinity of the electronic circulations. In the simplest approximation the current circulation is replaced by an equivalent point dipole at the centre of the ring²³ and for benzene this gives, for points in the plane of the ring,

$$\delta_{RC}(\text{benzene}) = -\frac{\chi_{RC}^{\perp}(\text{benzene})}{3R^3} \quad (5)$$

where R is the distance of the nucleus from the centre of the ring. For the hydrogen atoms in benzene this leads to a chemical shift of $\delta_{RC}(\text{benzene}) = -0.035\chi_{RC}^{\perp}(\text{benzene}) = 1.46$ p.p.m. Of course it is still necessary to appeal to model compounds in order to derive a chemical shift in the absence of ring-current effects (δ_{local}). Taking the local contribution as the experimental chemical shift of the olefinic protons in 1,3-cyclohexadiene gives $\delta_{\text{local}} = 5.77$ p.p.m. Thus the Pople model for benzene leads to $\delta(\text{benzene}) = \delta_{\text{local}} + \delta_{RC} = 7.23$ p.p.m., in good agreement with the experimental value of $\delta = 7.27$ p.p.m. In a magnetically isotropic molecule with the π -electrons confined to the surface of a sphere with radius r , the ring current contribution to the chemical shift of a central atom takes the form¹⁰

$$\delta_{RC}(C_{60}) = 2 \frac{\chi_{RC}(C_{60})}{r^3} \quad (6)$$

Because the ring-current magnetic susceptibility was found to be small in C_{60} , the chemical shift for an atom at the centre was also predicted to be small¹⁰.

The first evidence for ring currents in C_{60} , came from theoretical and experimental developments that provided information on the currents flowing along the bonds and around the individual rings in the presence of a magnetic field. Theoretical refinements to the London theory by Pasquarello *et al.*^{23,24} allowed the calculation of the bond currents, and proton NMR measurements on derivatives of C_{60} by Wudl, Diederich and Smith and their co-workers²⁵⁻³⁰ showed that the protons attached near the surface of the sphere were subject to well defined shielding effects that were originating from the C_{60} π -electrons. Because the fullerenes are continuous aromatic compounds, the placement of nuclei near the surface to act as probes of the secondary magnetic fields necessarily involves the attachment of substituents to the conjugated carbon atoms, which perturbs the electronic structure. Experimentalists therefore faced the challenge of developing a procedure that placed protons in a well defined relationship to the fullerene surface without

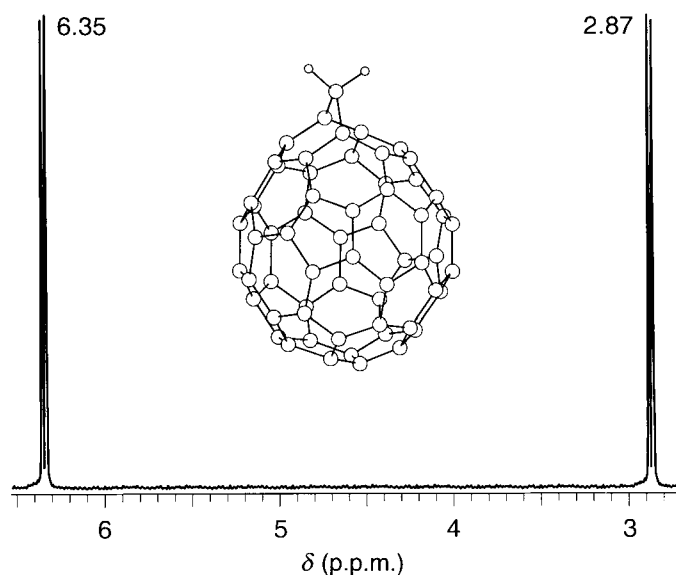


FIG. 3 Proton NMR spectrum of the fulleroid ($C_{61}H_2$). Redrafted from a figure provided by Wudl²⁵.

destroying the pattern of conjugation. The most dramatic application of this principle to date is represented by the fulleroids³⁰, in which a bridge group (for example the CH_2 unit in Fig. 3) is added across a fullerene bond. These derivatives retain the basic fullerene electronic structure by virtue of a strong homoaromatic interaction; thus, the fulleroids may be taken as the minimally perturbed fullerene. The proton NMR spectrum of the parent fulleroid is shown in Fig. 3, which demonstrates that the protons on the bridge group are subjected to entirely different magnetic environments²⁵. With chemical shifts of $\delta = 2.87$ and 6.35 p.p.m., the down-field proton in particular is experiencing a very strong secondary magnetic field—as would arise if it were positioned over a region of the molecule in which there is a strong paramagnetic ring current.

This result was in striking agreement with the pattern of π -electron ring currents calculated for C_{60} by Pasquarello *et al.*^{23,24} using an extension of the London theory. The calculations predicted the existence of large ring currents in C_{60} (Fig. 4), comparable in strength to those in benzenoid hydrocarbons. Of particular importance, however, was the finding of paramagnetic currents in the five-membered rings (5-MRs). These currents were found to be sufficiently large that their sum over the twelve 5-MRs almost exactly cancelled the diamagnetic currents in the twenty 6-MRs. Thus the cancellation was shown to be a function of the fullerene topology, with the diamagnetic term generally largest in the 6-MRs and the paramagnetic term predominant in the 5-MRs, in a proportion such that the overall effect of the π -electrons on the magnetic susceptibility was negligible. These findings seemingly rationalized all of the magnetic properties of the fullerenes, showing that the existence of currents in the individual rings (segregated ring currents)²⁵ in C_{60} was entirely consistent with the vanishingly small ring-current contribution to the total magnetic susceptibility of C_{60} , as calculated theoretically and measured experimentally.

The large difference in the magnetism of C_{60} and C_{70} served to dispel the idea that there would be a readily observable and characteristic fullerene ring-current effect, but the existence of ring currents, as revealed by the NMR data, demonstrated that the fullerenes possess the same sort of delocalized electronic structure as is found in normal aromatic organic molecules. Nevertheless, the signature of their electronic structure is modified in two important respects³¹, and they are best considered as species of ambiguous aromatic character¹⁴. First, the continuity of their electronic structure—the absence of any periphery or

boundary—does not allow the usual aromatic substitution reactions. Second, the enormous strain present in the fullerene structure predisposes the carbon atoms towards the tetrahedral geometry necessary for addition reactions. Although the fullerenes have shown a singular chemistry³², characteristic of an olefinic electronic structure, they are properly considered as a unique class of strained and continuous aromatic molecules³¹.

Endohedral fullerenes

The final piece of the picture of fullerene magnetism was supplied by Saunders and co-workers³³, who synthesized $^3He@$ fullerenes by pressurization experiments, where $X@C_{60}$ denotes an atom X within (endohedral to) C_{60} . By choosing the most inert element in the periodic table, as well as an ideal nucleus for NMR detection, they finally realized the experiment postulated by Kroto *et al.* in 1985 (ref. 8), providing a window to the interior environment of the fullerenes. The 3He NMR spectra of $^3He@C_{60}$ and $^3He@C_{70}$ are shown in Fig. 5—the endohedral nuclei are shielded by 6 and 29 p.p.m., respectively, with respect to dissolved 3He . Whereas a substantial chemical shift was expected for $^3He@C_{70}$ on the basis of the appreciable π -electron ring-current contribution to the magnetic susceptibility^{14,15}, the chemical shift ($\delta = -6$ p.p.m.) found for $^3He@C_{60}$ posed more of a problem for the prevailing picture of C_{60} magnetism, for as

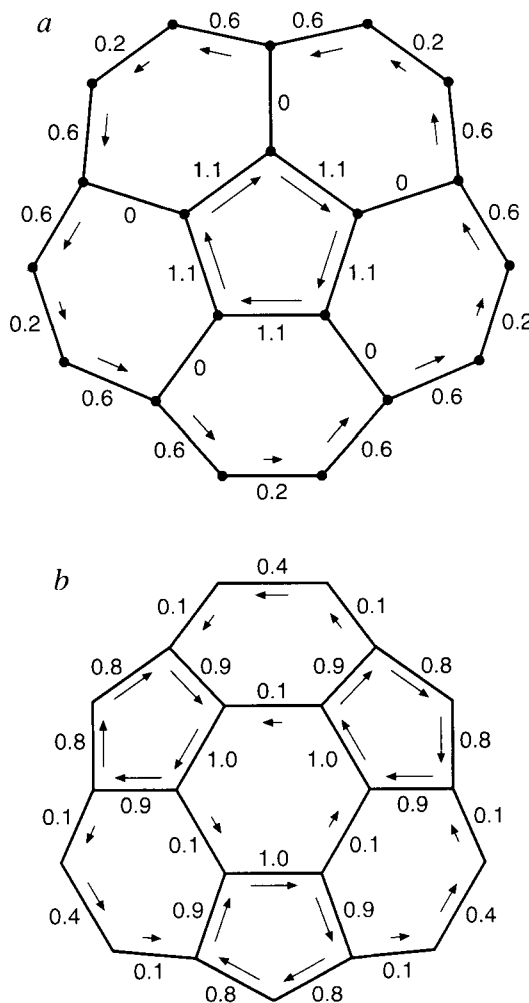


FIG. 4 π -electron ring currents in C_{60} for a magnetic field oriented perpendicular to: a, a five-membered ring (5-MR) and b, a six-membered ring (6-MR), calculated using the London theory. The currents in the 5-MRs are paramagnetic (clockwise), whereas those in the 6-MRs are mostly diamagnetic. The current strength is given with respect to that in benzene. From Pasquarello *et al.*^{23,24}.

equation (6) shows, the ring-current chemical shift experienced by a central atom is governed by the ring-current contribution to the *total* magnetic susceptibility¹⁰. This is quite unlike the situation in the functionalized fullerenes, where the nuclei observed in NMR experiments are sensitive to the *local* ring currents on the surface of the sphere (Figs 3 and 4). Thus the experiments by Saunders *et al.*³³ made it difficult to turn the outside in, as far as the magnetic properties of C₆₀ are concerned³⁴.

From equation (6)¹⁰, $\delta_{RC}(C_{60}) = 0.079\chi_{RC}(C_{60})$, indicating that the relationships of the ring-current chemical shift to the ring-current magnetic susceptibility in benzene (equation (5)) and C₆₀ are similar; the proportionalities are even closer if the rotationally averaged susceptibility is used for benzene, $\delta_{RC}(\text{benzene}) = -0.106\chi_{RC}^{av}(\text{benzene})$, although the sense of the relationship is opposite. Thus, to accommodate the chemical shift ($\delta = -6$ p.p.m.) found for ³He@C₆₀ requires a ring-current contribution to the magnetic susceptibility of $\chi_{RC}(C_{60}) \approx -80$ CGS p.p.m. mol⁻¹ $\approx 2.4\chi_{RC}(\text{benzene})$, towards the upper end of the original estimate of the C₆₀ ring-current diamagnetism (Fig. 2), and quite difficult to rationalize with the experimental measurements of the bulk (total) magnetic susceptibility ($\chi(C_{60}) \approx -260$ CGS p.p.m. mol⁻¹).

The resolution of this apparent contradiction came from the realization that the endohedral chemical shift in the ³He@fullerenes does not result solely from the shielding by π -electron ring currents. Beginning in 1994, *ab initio* calculations by Cioslowski^{35,36} and Bühl *et al.*³⁷ reproduced the chemical shift of ³He@C₆₀ and ³He@C₇₀ with a reasonable degree of accuracy, although the results were extremely sensitive to the fullerene geometry, in agreement with the original London calculations (Fig. 2). However, additional calculations by Bühl *et al.*³⁷, threw new light on the origin of the endohedral chemical shift in the ³He@fullerenes. They calculated endohedral chemical shifts of $\delta = -5.2$ and -5.4 p.p.m. for ³He@C₆₀H₆₀ and ³He@C₇₀H₇₀, respectively. The calculations on the fully saturated analogues provided the first evidence that the endohedral chemical shifts are not solely due to π -electron ring currents. If these values are adopted for the local component of endohedral fullerene shielding (δ_{local}), then it is clear that the σ -bond anisotropy makes the principal contribution to the endohedral chemical shift in C₆₀.

Further support for this view is provided in Fig. 6, which shows a plot of the chemical shifts for a number of higher fullerenes, as obtained from London and *ab initio* calculations (refs 38, 39 and M. Bühl, W. Thiel and R.C.H., unpublished results). There is a reasonable linear correlation between the two sets of calculations, but the straight-line fit does not pass through the origin. The intercept of the line with the vertical axis gives the constant term that is omitted in the London ring-current calculations, namely the local contribution (mainly arising from the σ -

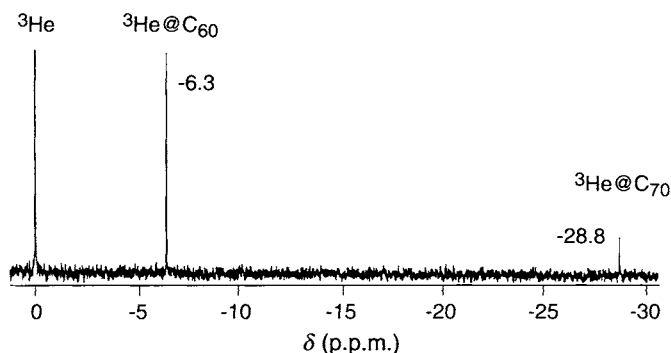


FIG. 5 The ³He NMR spectra of ³He@C₆₀ and ³He@C₇₀ together with dissolved ³He in 1-methylnaphthalene solution. Redrafted from a figure provided by Jiménez-Vázquez³³.

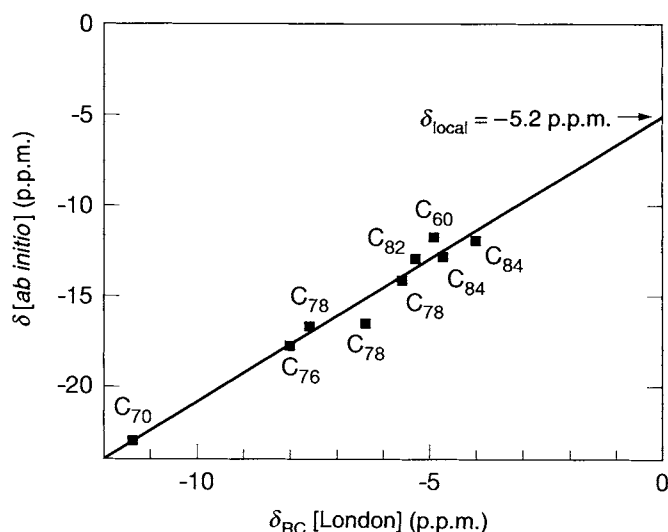


FIG. 6 Results of *ab initio* calculations for the total chemical shift versus London calculations of the ring-current-induced chemical shifts of a central atom in the higher fullerenes. Taken from Bühl and Thiel³⁹.

electrons), which can be taken to be roughly constant for the fullerenes studied: $\delta_{local} = -5.2$ p.p.m., in agreement with the estimates given above from calculations³⁷ on saturated endohedrals. The other important piece of information from this plot is the deviation of the slope from unity. Although it is clear that the two calculations are capturing the same phenomena, the London calculations underestimate the effectiveness of the ring currents in shielding the endohedral atom. Given that the *ab initio* calculations also underestimate the shielding in ³He@C₇₀ (Fig. 6), the London calculations give ring-current chemical shifts that are about half the values required to reproduce experimental values. By the same token the London calculations seem to reproduce the magnetic susceptibilities quite well, whereas current *ab initio* calculations overestimate this quantity^{37,39,40}. In light of these results, a ring-current chemical shift of $\delta_{RC}(C_{60}) \approx -1$ p.p.m., would be accommodated by $\chi_{RC}(C_{60}) \approx -10$ CGS p.p.m. mol⁻¹ $\approx 1/4\chi_{RC}(\text{benzene})$. Taken together, the data allow two important conclusions. First, the curvature in the fullerenes leads to additional shielding for an internal endohedral atom over that calculated from the London ring currents. This level of theory has been quite successful in planar conjugated systems^{41,44}, and it is apparent that details of the induced currents in the fullerenes (recently revealed in calculations by Zanasi and Fowler⁴⁵) serve to magnify their shielding effect on an endohedral nucleus without changing their contribution to the magnetic susceptibility. Second, it is clear that in the five years since the isolation of the fullerenes, it has become possible to present a rather complete picture of fullerene magnetism. To paraphrase Kroto *et al.*⁸, it might be said that C₆₀ ring currents make a remarkably small contribution to the chemical shift in the NMR of a central atom.

Thus, ring-current effects on the magnetism of C₆₀ are exceedingly subtle. Ring-current circulations occur on a length scale that encompasses the individual rings. Although there is near-cancellation of individual ring-current contributions to the susceptibility and to the chemical shift of a central atom, these ring currents influence the chemical shifts of protons attached near the surface. Recently, careful magnetic measurements by Hermans, Luo, Ramirez and their co-workers^{46,48} have found evidence for a small change ($\sim 1\%$) in the magnetic susceptibility of C₆₀ due to solid-state effects. Previously, neutron-scattering experiments by David *et al.*⁴⁹ provided evidence for a change in the structure of the C₆₀ molecule below the orientational ordering transition at 259 K. The small jump in magnetic sus-

ceptibility seen in solid C_{60} at this temperature (Fig. 1 inset), can be accommodated by a variation in the ring-current contribution to the magnetic susceptibility as a result of a relative change in C_{60} bond lengths of about 0.003 Å (ref. 48; see Fig. 2).

Giant fullerenes, nanotubes and graphite

With the large increase in diamagnetism and endohedral chemical shift on progressing from C_{60} to C_{70} , there has been speculation on the manner in which this trend would continue with the higher fullerenes. With the number of (paramagnetic) 5-MRs fixed at 12, and the fraction of (diamagnetic) 6-MRs increasing with fullerene size, it might be expected that ring-current effects would continue to increase in the higher fullerenes. Recent theoretical and experimental work has now shown that this does not occur in the known fullerenes^{38,39,50}. All of the currently available higher fullerenes show ^3He endohedral chemical shifts that fall between C_{60} and C_{70} (Fig. 6). The only known species so far predicted to resonate at higher field than $^3\text{He}@C_{70}$ is $^3\text{He}@C_{60}^-$, for which δ values of -29 to -58 p.p.m. have been calculated^{10,35,37}. The ^{13}C NMR spectrum of a solution apparently containing C_{60}^- shows a remarkable deshielding with respect to neutral C_{60} (ref. 51).

The variations seen in the magnetic susceptibility and endohedral chemical shift in low-symmetry fullerenes^{38,39,50} may be attributed to quantum size effects, as seen in model calculations on small metallic clusters⁵². These apparently random variations occur as a result of fluctuations in the pattern of eigenvalues and nodes in the wave function due to the topology of the cluster^{38,39}. Recent London calculations³⁸ of the magnetic properties of the giant icosahedral fullerenes have demonstrated the beginnings of a trend towards the very large diamagnetism of graphite (Fig. 7). In particular, C_{540} is predicted to exhibit more than 10% of the π -electron magnetic susceptibility of graphite on a per carbon basis (Fig. 7). The endohedral chemical shifts are predicted to be invariant to cluster size, but subject to the quantum size effects seen in smaller fullerenes.

The previous theoretical treatments have been applied to molecules of finite size. At the other end of the spectrum are band-structure calculations, in which some or all of the boundary conditions are taken to infinity. The first successful quantum-mechanical treatment of the magnetic susceptibility of the graphite π -electrons was carried out by McClure^{53,54}. He was able to reproduce the high-temperature experimental data by use of a two-dimensional band-structure calculation on a single graphite sheet. The treatment is equivalent to the London theory, although the semimetallic character of graphite introduces some

new features. Because the conduction and valence bands in graphite touch at the corners of the hexagonal Brillouin zone, the interactions between the states near the Fermi level must be included. As a result the two-dimensional π -electron magnetic susceptibility of graphite is proportional to β^2 , whereas that of small molecules is linearly dependent on β , the resonance integral between neighbouring p_π orbitals.

Carbon compounds with length scales between those of the fullerenes and graphite have also been studied. London calculations on condensed ring systems with infinite length but finite width are able to explain the susceptibilities of pre-graphitic carbons in the size regime of crystallites (~ 100 Å)^{55,56}. The carbon nanotubes^{57,58} represent a further variation on the theme of curved carbon structures with a strong relationship to graphite. Recently Lu⁵⁹ adapted the London formalism to carry out band-structure calculations of the magnetic susceptibility of nanotubes, and found good agreement with the experimental results^{48,60}. In this treatment the nanotubes are represented as infinite cylinders of arbitrary radius. As in graphite, the energy gap in the nanotubes is small and a magnetic field exerts important effects on the electronic structure. Semiquantitative agreement with experiment^{48,60} (Fig. 1) is found for a nanotube of radius ~ 100 Å. Recent experiments by Chauvet and co-workers⁶¹ have demonstrated the presence of a significant degree of magnetic anisotropy in aligned nanotubes, with the largest diamagnetism occurring when the field is directed along the nanotube axis. This behaviour may be ascribed to the possibility of extended ring-current circulations around the nanotube axis⁴⁸.

Thus, despite its severe approximations, the London theory of ring currents⁶² provides an adequate theoretical account of all of the conjugated forms of carbon—large and small, curved and flat. The only adjustable parameter is the resonance integral β , and one measure of the success of this theory may be given by noting that in all of these different magnetic studies the authors were able to obtain reasonable agreement with experiment with $\beta = 2\text{--}5$ eV. Although it is difficult to prove the ring-current model with detailed calculations⁶³, its usefulness is beyond question.

It is now clear that C_{60} is not a sphere with π -electrons free to precess in a magnetic field. Yet studies of fullerene magnetism show that there are mobile π -electrons that respond to the field, but because they are constrained to move along the bonds of a polyhedral structure, their response is often quenched by a large Van Vleck paramagnetism. This feature, together with the resulting paramagnetic ring currents in the 5-MRs, leads to ambiguities in the conventional interpretation of fullerene magnetism in terms of aromatic character. In the final analysis C_{60} con-

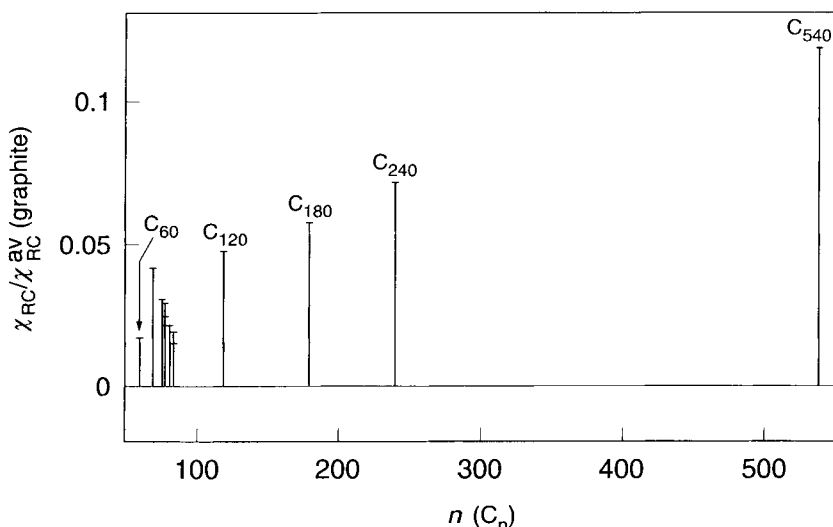


FIG. 7 Ring-current magnetic susceptibilities (relative to graphite on a per carbon basis, see Fig. 1), as a function of fullerene mass, calculated with the London theory. From Haddon and Pasquarello³⁸.

spires to possess magnetic properties common to both of the other carbon allotropes: it has a feeble diamagnetism, much like diamond (Fig. 1), but possesses ring currents, as in graphite. □

R. C. Haddon is at AT&T Laboratories, Murray Hill, New Jersey 07974-0636, USA.

1. Raman, C. V. *Nature* **123**, 945 (1929).
2. Ehrenfest, P. *Physica* **5**, 388–391 (1925).
3. Pauling, L. *J. chem. Phys.* **4**, 673–677 (1936).
4. Lonsdale, K. *Proc. R. Soc.* **159**, 149–161 (1937).
5. Van Vleck, J. H. *The Theory of Electric and Magnetic Susceptibilities* (Oxford Univ. Press, 1965).
6. Garratt, P. J. *Aromaticity* (Wiley, New York, 1986).
7. Minkin, V. I., Glukhovtsev, J. N. & Simkin, B. Y. *Aromaticity and Antiaromaticity* (Wiley, New York, 1994).
8. Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. *Nature* **318**, 162–164 (1985).
9. Smalley, R. E. *Chem. Engng News* **66**(35), 33–35, Aug. 29 (1988).
10. Elser, V. & Haddon, R. C. *Nature* **325**, 792–794 (1987).
11. Elser, V. & Haddon, R. C. *Phys. Rev.* **A36**, 4579–4584 (1987).
12. Haign, C. W. & Mallon, R. B. *Prog. NMR Spectrosc.* **13**, 303–344 (1980).
13. Krätschmer, W., Lamb, L. D., Fostiropoulos, K. & Huffman, D. R. *Nature* **347**, 354–358 (1990).
14. Haddon, R. C. et al. *Nature* **350**, 46–47 (1991).
15. Ruoff, R. S. et al. *J. phys. Chem.* **95**, 3457–3459 (1991).
16. Hückel, E. *Z. Phys.* **70**, 204–286 (1931).
17. Hückel, E. *Z. Phys.* **60**, 423–456 (1930).
18. Haddon, R. C. *Acc. chem. Res.* **21**, 243–249 (1988).
19. Dauben, H. J., Wilson, J. D. & Laity, J. L. *J. Am. chem. Soc.* **91**, 1991–1998 (1969).
20. Haddon, R. C. & Elser, V. *Chem. Phys. Lett.* **169**, 362–364 (1990).
21. Schmaltz, T. G. *Chem. Phys. Lett.* **175**, 3–5 (1990).
22. Pople, J. A. *J. chem. Phys.* **24**, 1111 (1956).
23. Pasquarello, A., Schluter, M. & Haddon, R. C. *Science* **257**, 1660–1661 (1992).
24. Pasquarello, A., Schluter, M. & Haddon, R. C. *Phys. Rev.* **A47**, 1783–1789 (1993).
25. Suzuki, T., Li, Q., Khemani, K. C. & Wudl, F. *J. Am. chem. Soc.* **114**, 7301–7302 (1992).
26. Prato, M., Suzuki, T., Wudl, F., Lucchini, V. & Maggini, M. *J. Am. chem. Soc.* **115**, 7876–7877 (1993).

27. Prato, M. et al. *J. Am. chem. Soc.* **115**, 8479–8480 (1993).
28. Isaacs, L., Wehrsig, A. & Diederich, F. *Helv. chim. Acta* **76**, 1231–1250 (1993).
29. Smith, A. B. et al. *J. chem. Soc., chem. Commun.* 2187–2188 (1994).
30. Wudl, F. *Acc. Chem. Res.* **25**, 157–161 (1992).
31. Haddon, R. C. *Science* **261**, 1545–1550 (1993).
32. Taylor, R. & Walton, D. R. M. *Nature* **363**, 685–693 (1993).
33. Saunders, M. et al. *Nature* **367**, 256–258 (1994).
34. Haddon, R. C. *Nature* **367**, 214 (1994).
35. Cioslowski, J. *J. Am. chem. Soc.* **116**, 3619–3620 (1994).
36. Cioslowski, J. *Chem. Phys. Lett.* **227**, 361–364 (1994).
37. Bühl, M. et al. *J. Am. chem. Soc.* **116**, 6005–6006 (1994).
38. Haddon, R. C. & Pasquarello, A. *Phys. Rev.* **B50**, 16459–16463 (1994).
39. Bühl, M. & Thiel, W. *Chem. Phys. Lett.* **233**, 585–589 (1995).
40. Fowler, P. W., Lazzeretti, P. & Zanasi, R. *Chem. Phys. Lett.* **165**, 79–86 (1990).
41. McWeeny, R. *Molec. Phys.* **1**, 311–321 (1958).
42. Mallion, R. B. *Molec. Phys.* **25**, 1415–1432 (1973).
43. Mallion, R. B. *J. chem. Phys.* **75**, 793–797 (1981).
44. Haddon, R. C. *Tetrahedron* **28**, 3613–3634, 3635–3655 (1972).
45. Zanasi, R. & Fowler, P. W. *Chem. Phys. Lett.* **238**, 270–280 (1995).
46. Heremans, J., Olk, C. H. & Morelli, D. T. *Phys. Rev.* **B49**, 15122–15125 (1994).
47. Luo, W., Wang, H., Ruoff, R. S., Cioslowski, J. & Phelps, S. *Phys. Rev. Lett.* **73**, 186–188 (1994).
48. Ramirez, A. P. et al. *Science* **265**, 84–86 (1994).
49. David, W. I. F. et al. *Nature* **353**, 147–149 (1991).
50. Saunders, M. et al. *J. Am. chem. Soc.* **117**, 9305–9308 (1995).
51. Bausch, J. W. et al. *J. Am. chem. Soc.* **113**, 3205–3206 (1991).
52. van Ruitenbeek, J. M. & van Leeuwen, D. A. *Phys. Rev. Lett.* **67**, 640–643 (1991).
53. McClure, J. W. *Phys. Rev.* **104**, 666–671 (1956).
54. McClure, J. W. *J. chim. Phys.* **57**, 859–865 (1960).
55. Hoarau, J. & Volpilhac, G. *Phys. Rev.* **B14**, 4045–4053 (1976).
56. Volpilhac, G. & Hoarau, J. *Phys. Rev.* **B17**, 1445–1449 (1978).
57. Iijima, S. *Nature* **354**, 56–58 (1991).
58. Ebbesen, T. W. & Ajayan, P. M. *Nature* **358**, 220–222 (1992).
59. Lu, J. P. *Phys. Rev. Lett.* **74**, 1123–1126 (1995).
60. Wang, X. K., Chang, R. P. H., Patashinski, A. & Ketterson, J. B. *J. Mater. Res.* **9**, 1578–1582 (1994).
61. Chauvet, O. et al. *Phys. Rev. B* (in the press).
62. London, F. *J. Phys. Radium, Paris* **8**, 397–409 (1937).
63. Fleischer, U., Kutzelnigg, W., Lazzeretti, P. & Mülenkamp, V. *J. Am. chem. Soc.* **116**, 5298–5306 (1994).

ACKNOWLEDGEMENTS. I thank V. Elser, A. Pasquarello, A. P. Ramirez and the late M. A. Schluter for their contributions to this work.

LETTERS TO NATURE

Evidence for shock acceleration of high-energy electrons in the supernova remnant SN1006

K. Koyama*, R. Petre†, E. V. Gotthelf†‡, U. Hwang†, M. Matsuura§, M. Ozaki* & S. S. Holt†

* Department of Physics, Kyoto University, Sakyo-ku, Kyoto 606-01, Japan

† Laboratory for High Energy Astrophysics, NASA/GSFC, Greenbelt, Maryland 20771, USA

‡ Universities Space Research Association, 7501 Forbes Blvd, Seabrook, Maryland 20706, USA

§ Department of Physics, Kyoto Sangyou University, Kita-ku, Kyoto 603, Japan

HIGH-ENERGY cosmic rays (relativistic heavy nuclei) play an important role in heating interstellar matter in the Milky Way^{1,2}, and they affect chemical abundances through collisions with atoms in the interstellar gas². Although it has long been thought that these cosmic rays arise from supernovae^{3,4}, direct evidence for such an association has been lacking. Here we report X-ray observations of the remnant of supernova 1006, made by the ASCA satellite, which indicate that emission from the edges of the remnant shell is dominated by radiation from electrons accelerated to energies of ~100 TeV within the shock front. Ions in the shell are likely to have been accelerated to similar energies, thus giving rise to very-high-energy cosmic rays.

The X-ray CCD cameras on the ASCA observatory (the Solid-state Imaging Spectrometers) offer unprecedented capability for

carrying out spatially resolved X-ray spectroscopy, with a broad bandpass (0.5–10 keV) and good spectral (~100 eV at 1 keV) and spatial (1 arcmin full width at half maximum) resolution⁵. Two 25-ks pointings were performed at the ~30-arcmin-diameter supernova remnant SN1006 with the 22 × 22 arcmin field of view CCD instruments: one centred on the remnant and the other on the northeastern rim. Figure 1 shows a mosaic 0.4–8.0 keV image of SN1006. As with the previous observations^{6,7}, we observe a shell-like structure, dominated by the northeastern (NE) and southwestern (SW) rims, whose surface brightness is an order of magnitude higher than in the other regions. The limb-brightened X-ray morphology is typical of supernova remnants (SNR) dominated by thermal processes in the shock-heated interstellar material and/or supernova ejecta.

Figure 2 shows X-ray spectra from regions along the bright NE rim (Fig. 2a) and near the (projected) centre of the remnant (Fig. 2b). The spectrum of the rim (Fig. 2a) is virtually featureless. The overall 0.4–8.0 keV spectrum is well fitted ($\chi^2 = 344$, with 332 degrees of freedom) by a model consisting of a power law with energy spectral index $\alpha = 1.95 \pm 0.20$, and two emission lines (K-shell resonance transitions of He-like and H-like O at 574 eV and 653 eV). Substitution of a thermal Bremsstrahlung model for the power law yields an unacceptable fit ($\chi^2 = 930$). As the bright rims are responsible for ~75% of the X-ray flux from SN1006 above 1 keV, this is the integrated spectrum measured by previous detectors lacking spatial resolution^{8–11}. The best-fit energy spectral index found by ASCA is steeper than the $\alpha = 1.2$ inferred from the earliest observations⁸, but consistent with the most sensitive observations¹¹. It is also steeper than the $\alpha \approx 1.0$ characteristic of the Crab Nebula and similar SNR, whose centrally concentrated X-ray emission arises from synchrotron radiation produced by electrons accelerated to high energy in the pulsar magnetosphere¹². The featureless spectrum from SN1006 is a

long-standing mystery. The X-ray spectra of all other young SNR with shell-like morphologies (for example, Cassiopeia A¹³⁻¹⁵, Tycho^{14,16} and Kepler¹⁷) are dominated by emission lines characteristic of shock-heated, optically thin plasma.

The important new information about SN1006 provided by ASCA is the spectral character of the remainder of the SNR. As shown in Fig. 2b, the spectrum of the interior contrasts sharply with that of the northwestern (NE) rim. It is dominated by emission lines characteristic of highly ionized elements, like the X-ray spectra of other young SNR. To our knowledge, this is the first detection of X-ray emission lines other than oxygen^{18,19} from SN1006, and their presence is highly suggestive of an origin in a hot plasma. The spectrum has been fitted using time-dependent ionization models from the hydrodynamic simulation codes (also known as non-equilibrium ionization models) of Hughes and Helfand²⁰ and Hamilton *et al.*²¹. The best-fit model ($\chi^2 = 251$, with 235 degrees of freedom), has an emission-weighted electron temperature of 1.8×10^7 K, and ionization timescale $nt \approx 300 \text{ cm}^{-3} \text{ yr}$ (where n and t are the electron density and elapsed time after the plasma was shock-heated to temperature kT). Although near-solar abundances of O, Ne and Fe are required, the Si abundance is at least an order of magnitude in excess of the solar value, and Mg and S relative abundances comparable to Si are allowed. The relative abundances are consistent with the calculated nucleosynthesis yields for type Ia supernovae (a supernova caused by accretion-induced ignition of a white dwarf, as opposed to core collapse of a massive star)²². The relatively low iron abundance is also consistent, as ultra-violet absorption line measurements show iron concentrated interior to the reverse shock and therefore not yet highly ionized²³⁻²⁵. A contribution from a hard, featureless component is required to obtain acceptable fits above 2 keV. The low nt value implies that the plasma is at least as far from ionization equilibrium as the younger remnants Tycho and Kepler²⁶. We

found no evidence for strong spatial variation of the spectrum with radius inside the shell or along the northwestern rim; that is, the spectrum shown in Fig. 2b is representative of everywhere in the remnant *except* the bright NE and SW rims. Also, the surface brightness of the oxygen lines is approximately the same at the rims as in the interior. This suggests that thermal emission pervades SN1006, but is dominated by a second, featureless component along the two rims.

Although previous measurements, most notably using the Imaging Proportional Counter on the Einstein satellite⁶, indicated strong spectral differences between the bright rims and elsewhere in the remnant, the ASCA data dramatically reveal the different character of the spectra from the two regions. Rosat High Resolution Imager observations reveal that in addition to this spectral difference there is a morphological difference: only for the bright segments is there a strong correlation between X-ray and radio surface brightness⁷.

Previous modelling efforts have focused on explaining the featureless continuum which dominates the integrated spectrum of SN1006 above 1 keV (refs 27–30). The generally accepted interpretation has been a shocked plasma under extreme non-equilibrium ionization conditions. This model, developed by Hamilton *et al.*^{28,29}, is based on a one-dimensional simulation of layered ejecta propagating into a uniform interstellar medium (ISM), wherein the X-ray emission arises primarily from reverse-shocked ejecta. Bremsstrahlung emission from fully ionized carbon in the outer ejecta layer gives rise to a power-law continuum X-ray spectrum with $\alpha = 1$ in this model. The extreme weakness (that is, absence) of the X-ray lines compared with SNRs like Tycho is simply a consequence of a weaker reverse shock propagating into the ejecta, which in turn is due to expansion of the forward shock into a substantially lower-ambient-density medium. As a result, the layer containing the intermediate-mass ejecta (Mg, Si, S) has not yet become sufficiently ionized to produce copious X-ray line emission.

The model of Hamilton *et al.*^{28,29}, which assumes that the spectrum is everywhere the same, cannot account for the observed spectral variations in SN1006, nor can it account for the revised value of α . The spectral differences along the rim require an azimuthal asymmetry in either the ISM or the ejecta, whereas a spectral index steeper than unity requires that the initial density of either the ISM or the outer ejecta of ionized carbon depends on the distance from the explosion centre. Whereas the spatial uniformity of the line emission suggests that oxygen and the intermediate-mass elements (Mg and Si in particular) are symmetrically distributed, the strong featureless continuum requires the fully ionized carbon to be concentrated in the bright rims. As C and O are intimately related in nucleosynthesis processes, no supernova model predicts different spatial distributions for these elements. Additionally, the strengths of the lines from the intermediate-mass elements, which arise deep within the ejecta layer, require a stronger reverse shock (and thus a generally denser ISM) than is allowed by the model of Hamilton *et al.* (the measured surface brightness of the Si line at 1.86 keV is a factor of five higher than predicted). Finally, as the volume emissivity of a thermal plasma is proportional to the square of the density, one would expect that the density of the X-ray-emitting plasma at the rims would be a factor of three higher than elsewhere from the relative surface brightness across the remnant. If this were so, then the line emission should be stronger relative to the continuum at the bright rims, and not weaker. These inconsistencies force us to conclude that in contrast with other young SNR, any thermal model attempting to describe SN1006 must be extremely complicated, requiring *ad hoc* assumptions about strong asymmetries in the ISM and possibly the ejecta.

A natural alternative is that the excess rim emission is intrinsically non-thermal^{27,30}, arising as synchrotron emission from electrons accelerated to relativistic energies within the shock region by the first-order Fermi mechanism³¹. This additional

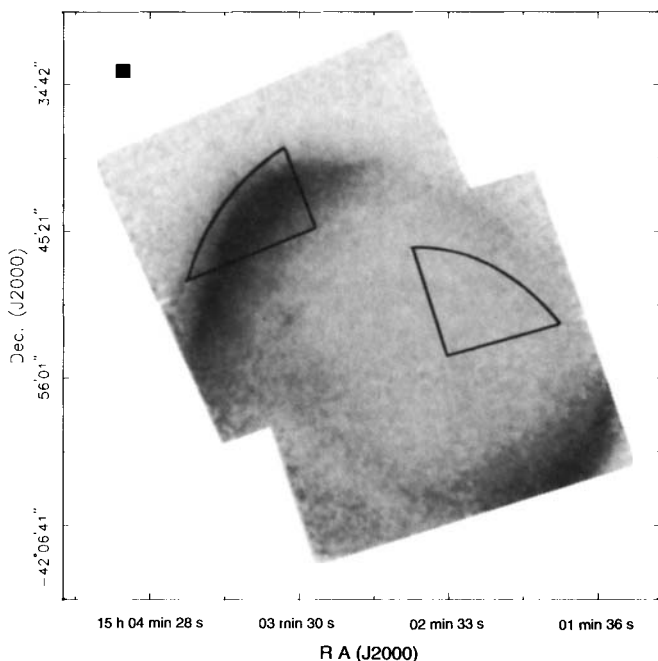
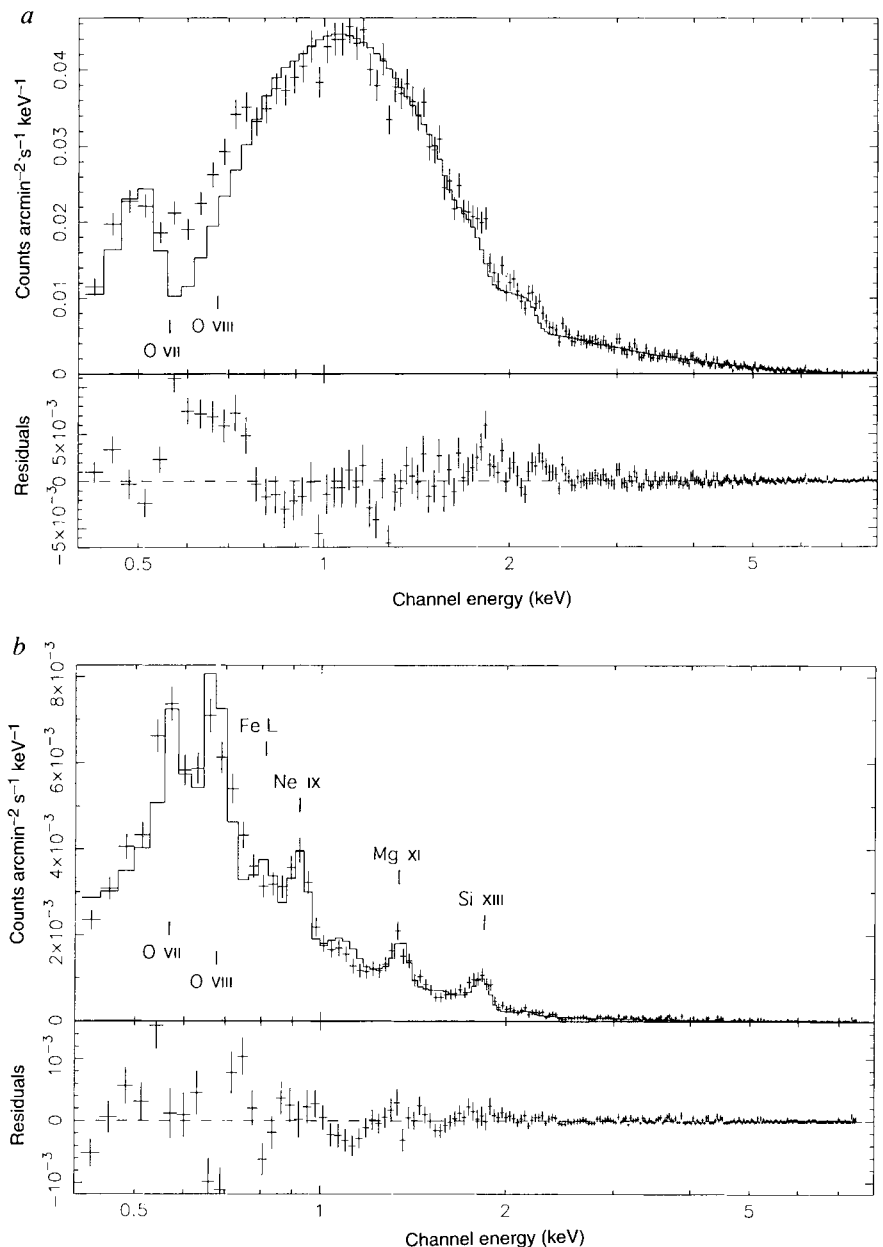


FIG. 1. CCD (charge-coupled-device) image of SN1006 in the 0.4–8 keV band, from the ASCA satellite. This image is an exposure-corrected mosaic of two pointings, one centred on the SNR and the other on the NE rim. The 'boxes' indicate the regions from which the spectra displayed in Fig. 2 were extracted. The odd shapes result from restricting the extraction to events from single CCD chips from the 4-CCD mosaicing comprising the Solid-state Imaging Spectrometer array for reduction of calibration uncertainties. The filled rectangle at the upper left is ~ 1 arcmin on a side.

FIG. 2 Spectrum of the bright NE rim of SN1006. The best-fit model consists of a power law with energy spectral index $\alpha = 1.95 \pm 0.20$, two narrow O lines at 0.574 keV and 0.653 keV, and a column density slightly in excess of Galactic ($N_H = 1.8 \times 10^{21} \text{ cm}^{-2}$). The solid line superposed on the data is the power-law model, without the O lines. The O lines appear blended in the plot of the residuals. *b*, Spectrum of the interior of SN1006. Narrow emission lines, as marked on the figure, are detected at energies of 0.574, 0.653, 0.915, 1.34 and 1.86 keV, corresponding to the K-shell resonance transitions of He- and H-like oxygen, and He-like Ne, Mg and S, respectively. The line at 0.84 keV corresponds to an L-shell transition of Fe. The solid curve represents the best-fit non-equilibrium ionization model plus a hard component.



component is superposed on a thermal spectrum resembling that of other young SNR, which can be described within the framework of the Hamilton *et al.* model^{28,29}. The synchrotron origin of the excess emission is supported by the strong correlation between the radio and X-ray surface brightness along the NE rim, in contrast to the lack of correlation along the thermal NW rim¹⁸. Published models of this process reproduce the flat radio spectrum and the steeper X-ray spectrum by predicting a power-law spectrum whose slope breaks between the radio and X-ray bands as synchrotron losses become important^{27,30}. Although these models do not accurately produce the surface brightness correlation, more recent work does so³².

The implications of non-thermal X-ray emission associated with a supernova remnant shock are important. The energy of a synchrotron photon E_p is related to the electron energy E_e and the magnetic field B by:

$$E_p \approx 4 \text{ keV} \times (B/1 \text{ mG}) \times (E_e/10 \text{ TeV})^2$$

There is no measurement of the magnetic field strength for any type Ia SNR (of which SN1006 is thought to be an example^{33,34}). The best estimate³⁵ is 6–10 μG . As the X-ray spectrum remains

flat up to $\sim 20 \text{ keV}$ (refs 8, 11), we infer that electrons of energy $\geq 200 \text{ TeV}$ are being produced in the SN1006 shock. These energies are among the highest inferred for electrons in any astrophysical setting.

More fundamentally, the potential identification of the SN1006 shell as a production site for such high-energy electrons provides the first strong observational evidence that very-high-energy cosmic rays are produced in SNR shocks. Theoretical treatments of particle acceleration via first-order Fermi processes do not draw distinctions between electrons and ions: highly relativistic particles ($E \gg m_p c^2$) of either charge are accelerated as the result of interaction with turbulence in the shock-compressed magnetic field³¹. Thus although the X-ray data provide direct evidence for the presence of very-high-energy electrons, there exists ample theoretical support for the presence of ions of similar energy—in other words, cosmic-ray nuclei. An energy of 200 TeV approaches the ‘knee’ of the cosmic-ray electron spectrum ($\sim 1,000 \text{ TeV}$), the energy at which it dramatically steepens. Our observations thus directly support for the first time the long-held assumption^{36,37} that SNRs accelerate cosmic rays all the way up to the ‘knee’, and provides a means for discriminating

between competing models for the rate of energy gain during diffusive shock acceleration^{36,37}.

Although SN1006 is the only strongly supported example thus far of a SNR with non-thermal X-ray emission from the shell, there exist data suggesting this process occurs in other remnants as well. The next best example is Cas A, where the 4–8 keV continuum has an energy spectral index $\alpha = 1.7$ and a morphology highly correlated with that of the radio continuum, in sharp contrast with that of the X-ray line emission¹⁵. Other candidates are IC443, whose hard (5–20 keV) X-ray component can be fitted by a power law with $\alpha \approx 2$, and Tycho, for which theoretical estimates suggest that the continuum above ~ 4 keV might be dominated by non-thermal emission³⁰. Careful spatially resolved X-ray spectroscopic studies of these and other remnants must be carried out before the prevalence of cosmic-ray acceleration in SNR shocks can be assessed. $\square$

Received 28 July; accepted 5 October 1995.

1. Berezhinskii, V. S., Bulanov, S. V., Dogiel, V. A., Ginzburg, G. L. & Ptuskin, V. S. *Astrophysics of Cosmic Rays* (North Holland, Amsterdam, 1990).
2. Wefel, J. P. in *Cosmic Rays, Supernovae and the Interstellar Medium* (eds Shapiro, M. M., Silberberg, R. & Wefel, J. P.) 29–55 (Kluwer, Dordrecht, 1991).
3. Shklovskii, I. S. *Dokl. Akad. Nauk SSSR* **91**, 475–478 (1953).
4. Ginzburg, V. L. *Dokl. Akad. Nauk SSSR* **92**, 1133–1136 (1953).
5. Tanaka, Y., Inoue, H. & Holt, S. S. *Publ. astr. Soc. Japan* **46**, L37–L41 (1994).
6. Pye, J. P. et al. *Mon. Not. R. astr. Soc.* **194**, 569–582 (1982).
7. Winkler, P. F. & Long, K. S. in *The Soft X-Ray Cosmos* (eds Schlegel, E. M. & Petre, R.) 312–314 (Am. Inst. Phys., New York, 1994).
8. Becker, R. H., Szymkowiak, A. E., Boldt, E. A., Holt, S. S. & Serlemitsos, P. J. *Astrophys. J.* **240**, L33–L35 (1980).
9. Koyama, K., Szymkowiak, A. E., Boldt, E. A., Holt, S. S. & Serlemitsos, P. J. *Publ. astr. Soc. Japan* **39**, 437–445 (1987).
10. Leahy, D. A., Nousek, J. & Hamilton, A. J. S. *Astrophys. J.* **374**, 218–226 (1991).
11. Ozaki, M., Koyama, K., Ueno, S. & Yamauchi, S. *Publ. astr. Soc. Japan* **46**, 367–373 (1994).
12. Asaoka, I. & Koyama, K. *Publ. astr. Soc. Japan* **42**, 625–632 (1990).
13. Becker, R. H. et al. *Astrophys. J.* **234**, L73–L76 (1979).
14. Petre, R. et al. in *UV and X-Ray Spectroscopy of Laboratory and Astrophysical Plasmas* (eds Silver, E. & Kahn, S.) 424–433 (Cambridge Univ. Press, 1993).
15. Holt, S. S., Gotthelf, E. V., Tsunemi, H. & Negoro, H. *Publ. astr. Soc. Japan* **46**, L151–L155 (1994).
16. Becker, R. H. et al. *Astrophys. J.* **235**, L5–L8 (1980).
17. Becker, R. H., White, N. E., Boldt, E. A., Holt, S. S. & Serlemitsos, P. J. *Astrophys. J.* **237**, L77–L79 (1980).
18. Galas, C. M. F., Venkatesan, D. & Garmire, G. P. *Astrophys. Lett.* **22**, 103–108 (1982).
19. Vartanian, M. H., Lum, K. S. K. & Ku, W. H.-M. *Astrophys. J.* **288**, L5–L9 (1985).
20. Hughes, J. P. & Helfand, D. J. *Astrophys. J.* **291**, 544–560 (1985).
21. Hamilton, A. J. S., Sarazin, C. L. & Chevalier, R. A. *Astrophys. J. Suppl. Ser.* **51**, 115–148 (1983).
22. Nomoto, K., Thielemann, F.-K. & Yokoi, K. *Astrophys. J.* **286**, 644–658 (1986).
23. Wu, C.-C., Leventhal, M., Sarazin, C. L. & Gull, T. R. *Astrophys. J.* **269**, L5–L8 (1983).
24. Fesen, R. A., Wu, C.-C., Leventhal, M. & Hamilton, A. J. S. *Astrophys. J.* **327**, 164–177 (1988).
25. Hamilton, A. J. S. & Fesen, R. A. *Astrophys. J.* **327**, 178–196 (1988).
26. Tsunemi, H. in *New Horizon of X-Ray Astronomy—First Results from ASCA* (eds Makino, F. & Ohashi, T.) 81–93 (Universal Academy Press, Tokyo, 1994).
27. Reynolds, S. P. & Chevalier, R. A. *Astrophys. J.* **245**, 912–919 (1981).
28. Hamilton, A. J. S., Sarazin, C. L., Szymkowiak, A. E. & Vartanian, M. H. *Astrophys. J.* **297**, L5–L9 (1985).
29. Hamilton, A. J. S., Sarazin, C. L. & Szymkowiak, A. E. *Astrophys. J.* **300**, 698–712 (1986).
30. Ammosov, A. E., Ksenofontov, L. T., Nikolaev, V. S. & Petukov, S. I. *Astr. Lett.* **20**, 157–162 (1994).
31. Blandford, R. D. & Eichler, D. *Phys. Rep.* **154**, 1–74 (1987).
32. Reynolds, S. P. *Astrophys. J.* (submitted).
33. Minkowski, R. *Astr. J.* **71**, 371–373 (1966).
34. Stephenson, F. G., Clark, D. & Crawford, D. F. *Mon. Not. R. astr. Soc.* **180**, 567–584 (1977).
35. Reynolds, S. P. & Ellison, D. C. in *Proc. 22nd Int. Cosmic Ray Conf.* Vol. 2 404–407 (Dublin Inst. for Advanced Studies, Dublin, 1991).
36. Lagage, P. O. & Cesarsky, C. J. *Astr. Astrophys.* **125**, 249–257 (1983).
37. Jokipii, R. J. *Astrophys. J.* **313**, 842–846 (1987).
38. Wang, Z. R., Asaoka, I., Hayakawa, S. & Koyama, K. *Publ. astr. Soc. Japan* **42**, 303–308 (1992).

ACKNOWLEDGEMENTS. U.H. is a NAS/NRC Research Associate. We thank S. Reynolds for insightful discussions.

Quantum confinement and light emission in SiO₂/Si superlattices

Z. H. Lu, D. J. Lockwood & J.-M. Baribeau

Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

PHOTONIC devices are becoming increasingly important in information and communication technologies. But attempts to integrate photonics with silicon-based microelectronics are hampered by the fact that silicon has an indirect band gap, which prevents efficient electron–photon energy conversion. Light-emitting silicon-based materials have been made using band-structure engineering of SiGe and SiC alloys and Si/Ge superlattices, and by exploiting quantum-confinement effects in nanoscale particles and crystallites^{1–3}. The discovery^{4,5} that silicon can be etched electrochemically into a highly porous form that emits light with a high quantum yield has opened up the latter approach to intensive study^{6–12}. Here we report the fabrication, by molecular-beam epitaxy, of well-defined superlattices of silicon and SiO₂, which emit visible light through photoluminescence. We show that this light emission can be explained in terms of quantum confinement of electrons in the two-dimensional silicon layers. These superlattice structures are robust and compatible with standard silicon technology.

The Si/SiO₂ superlattices were grown at room temperature on phosphorus-doped n-type (100) Si wafers by molecular-beam epitaxy in a VG Semicon V80 system. Thin Si films of various thickness were deposited on the ultraviolet/ozone-oxidized Si(100) wafer surface. A SiO₂ film ~ 1 nm thick was then grown *ex situ* by a rate-limited ultraviolet/ozone oxidation process. This procedure was repeated until a six-period Si/SiO₂ superlattice was made. The Si layer thickness and density were determined by small-angle X-ray scattering¹³. Modelling of the

reflectivity profiles established the periodic nature of the structure and revealed the Si layers to be of the same high density as found in vacuum-deposited Si. This suggests that the concentration of hydrogen, if there is any, incorporated in the Si film during growth should be very low. The chemical modulation and high purity of the superlattices were confirmed by Ar⁺-sputter profiling using Auger electron spectroscopy in a PHI 650 system. The physical structure of the superlattices was also studied by cross-sectional transmission electron microscopy (TEM) using a Philips EM 430 TEM operated at 300 kV. Figure 1 shows a cross-sectional TEM micrograph taken on a six-period SiO₂/Si superlattice with Si layer thickness of 2.8 nm and SiO₂

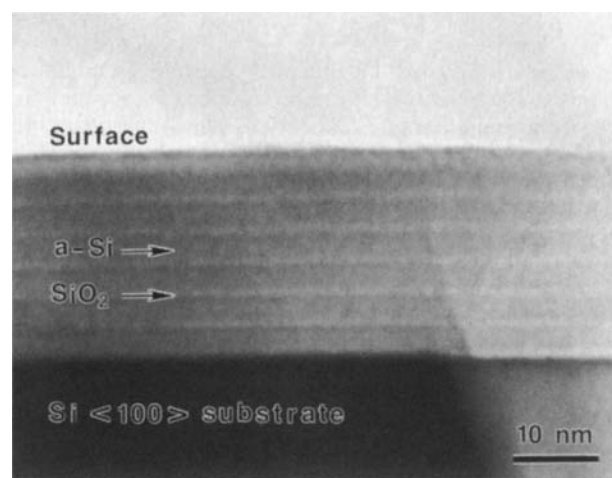


FIG. 1 Cross-sectional TEM micrograph taken from a six-period SiO₂/Si superlattice grown on Si(100) substrate. The alternating bright and dark bands represent SiO₂ and a-Si layers, respectively, as marked. The changes in the contrast on the right-hand side of the picture are caused by cleavage steps during TEM sample preparation.

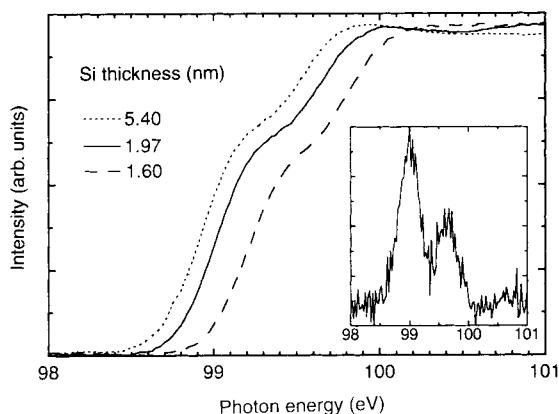


FIG. 2 Synchrotron radiation Si $L_{2,3}$ edge absorption spectra taken on SiO_2/Si superlattices with various Si film thicknesses as shown. Inset, derivative spectrum of the $L_{2,3}$ edge absorption from the SiO_2/Si superlattices with a Si film thickness of 1.97 nm.

film thickness of 1 nm. The SiO_2 and Si layers appear as white and dark bands, respectively; this contrast is caused by the large Si atomic density difference between the SiO_2 (2.2×10^{22} atoms cm^{-3}) and Si (5×10^{23} atoms cm^{-3}). The above TEM data revealed a uniform periodicity and abrupt (~ 0.5 nm) SiO_2/Si interfaces. The data also indicate that both the Si and SiO_2 layers are amorphous, as confirmed by Raman scattering measurements¹³.

One of the important effects of quantum confinement is the modification of the electronic energy band gap in these two-dimensional sandwiched Si films. The changes in the band gap can occur at both the conduction band minimum (CBM) and valence band maximum (VBM). To probe the shift in the CBM, we used soft X-ray Si $L_{2,3}$ edge absorption spectroscopy. The general principle of $L_{2,3}$ absorption spectroscopy is to excite the occupied core shell $p_{3/2,1/2}$ electrons to the empty conduction band. The absorption threshold is related to the CBM s -like electronic states due to the dipole selection rule. The absolute CBM position may not be determined accurately, but the relative CBM shift can be measured precisely. The experiments were carried out using the synchrotron radiation light source at the Synchrotron Radiation Center, Madison-Wisconsin. The highly monochromatic tunable soft X-ray beam near the Si $L_{2,3}$ edge was generated by the Canadian grasshopper beamline. The $L_{2,3}$ edge absorption was measured using a total electron yield method inside a vacuum chamber with a base pressure in the low 10^{-8} torr range. Figure 2 shows three representative experimental data sets obtained from different superlattices with Si film thickness of 5.40 nm, 1.97 nm and 1.60 nm, respectively. The data show clearly that the $L_{2,3}$ absorption edge increases in energy with decreasing Si film thickness. It is known¹⁴ that the $p_{3/2,1/2}$ energy is determined by the number of valence electrons. In the present case, the Si atoms in all films are the same and all have an average of four nearest-neighbour Si atoms, that is, an average of four valence electrons per Si atom. The changes in the $L_{2,3}$ edge are therefore caused by the shift in the CBM. To determine the relative shift, we took a first-order derivative of the spectrum (see Fig. 2 inset). The doublet peaks are the spin-orbit split $p_{3/2}$ and $p_{1/2}$ components of the L-shell electrons. The $p_{3/2}$ peak position at ~ 99 eV can then be determined precisely by a least-squares curve-fitting method. The relative CBM shift, ΔE_{CBM} , can then be determined by the energy difference of the $p_{3/2}$ peaks relative to that of a thick (>10 -nm) Si film where bulk band structure can be assumed. Figure 3 shows the CBM shift as a function of Si film thickness. The filled circles are the experimental data and the solid line is a least-squares fit to the data. It is found that the data are well represented by $\Delta E_{\text{CBM}} = 0.7/d_{\text{Si}}^2$, where d_{Si} is the Si film thickness.

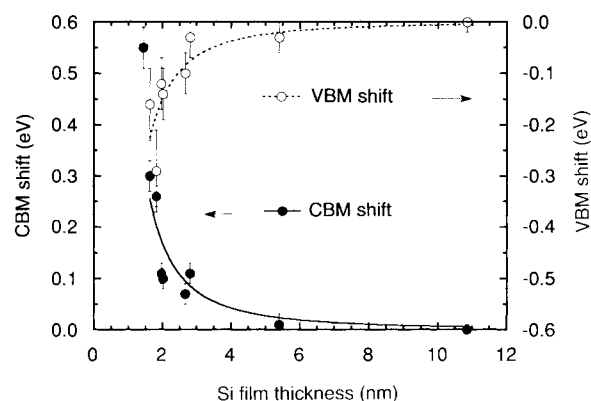


FIG. 3 Shifts in conduction band minimum (CBM) and valence band maximum (VBM) as a function of Si film thickness. The CBM was determined by synchrotron radiation Si $L_{2,3}$ edge absorption and VBM was determined by X-ray photoelectron spectroscopy. Filled circles are the experimental CBM shift data and the solid line is the least-squares fit. The open circles are the experimental VBM shift data and the dotted line is the least-squares fit.

The occupied valence band density of states was measured by X-ray photoelectron spectroscopy (XPS) using a PHI 5500 system, which is equipped with a monochromated Al $K\alpha$ source and a hemispherical electron analyser. A 75° photoelectron take-off angle was used to enhance the bulk sensitivity. To obtain high-resolution spectra, all data were taken at a pass energy of 5.75 eV. The valence band maximum was determined using the methodology described in ref. 15. Silicon $2p$ core levels were also recorded so that any possible artefacts such as surface charging and band bending could be eliminated. The XPS data are also shown in Fig. 3, where the zero energy shift is with reference to the thick film where bulk band structure is assumed. The open circles are the experimental data and the dotted line is the least-squares fit to the data. The data are fitted by $\Delta E_{\text{VBM}} (\text{eV}) = -0.5/d_{\text{Si}}^2$.

According to effective mass theory¹⁶ and assuming an infinite potential barrier at the interfaces, which should be a reasonable approximation because the band off-set at the SiO_2/Si interface is several electron volts (ref. 17), the energy gap variation, ΔE , for one-dimensionally confined Si should be determined by $\Delta E = \Delta E_{\text{CBM}} + \Delta E_{\text{VBM}} = (\alpha + \beta)/d_{\text{Si}}^2$, where α and β are constants determined by the electron and hole effective masses, respectively. This theory is in good agreement with the above experimental results. In more sophisticated calculations for nanoscale crystalline Si clusters^{9,12}, the blue shift of the crystal band gap

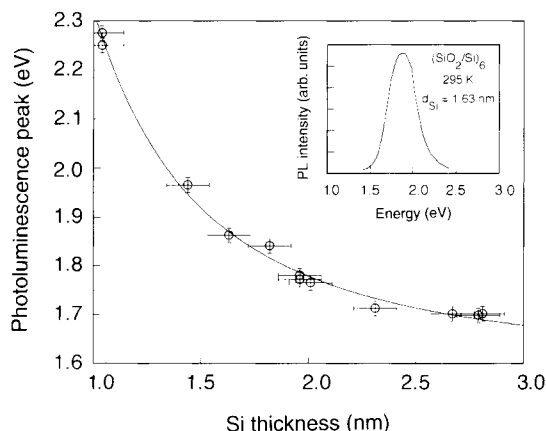


FIG. 4 Photoluminescence (PL) peak energy as a function of Si film thickness. The open circles are the experimental data and the solid line is the least-squares fit to the data. Inset, a raw PL spectrum taken from a superlattice with Si film thickness of 1.63 nm.

has been found to scale with crystal size, L , approximately as $1/L^\gamma$, with γ in the range 1.2–1.8. These calculations are consistent with our quantum-confinement interpretation.

The room-temperature photoluminescence (PL) properties of these superlattices was measured by using an 80-mW 457.9-nm argon laser beam as the excitation source, and by using a Photometrics CCD 9000 Si array detector to detect the luminescence which was dispersed by a Spex 1877C triple spectrometer. A PL peak at wavelengths across the visible range was observed from superlattices with Si film thicknesses ≤ 3 nm. A typical PL spectrum is shown in Fig. 4 inset. We could not detect any PL signal from superlattices with Si film thickness > 3 nm. This may be explained as due to a dramatic reduction in the quantum confinement effects, as shown in Fig. 3. The PL peak energy as a function of the Si thickness is plotted in Fig. 4. The peak energy position, E_{PL} , was found to be best-fitted by E_{PL} (eV) = $1.6 + 0.7/d_{\text{Si}}^2$. According to effective mass quantum-confinement theory, 1.6 eV corresponds to the band gap of bulk material. This fitted value is certainly larger than that of crystalline silicon, but is in good agreement with that of bulk non-crystalline silicon (1.5–1.6 eV at 295 K)¹⁸. Thus the PL data further confirm that quantum confinement governs the Si energy band structure and is the fundamental source of the bright Si luminescence. It should

be possible to tune the wavelength of the emitted light by varying the Si film thickness, which can be fabricated with atomic-layer precision by modern molecular-beam epitaxy technology. $\square$

Received 30 August; accepted 6 October 1995.

1. Iyer, S. S. & Xie, Y.-H. *Science* **260**, 40–46 (1993).
2. Abeles, B. & Tiedje, T. *Phys. Rev. Lett.* **51**, 2003–2006 (1983).
3. Wilson, W. L., Szajowski, P. F. & Brus, L. E. *Science* **262**, 1242–1244 (1993).
4. Canham, L. T. *Appl. Phys. Lett.* **57**, 1046–1048 (1990).
5. Cullis, A. G. & Canham, L. T. *Nature* **353**, 335–338 (1991).
6. Lehmann, V. & Gösele, U. *Appl. Phys. Lett.* **58**, 856–858 (1991).
7. Halimaoui, A. et al. *Appl. Phys. Lett.* **59**, 304–306 (1991).
8. Petrova-Koch, V. et al. *Appl. Phys. Lett.* **61**, 943–945 (1992).
9. Proot, J. P., Delerue, C. & Allan, G. *Appl. Phys. Lett.* **61**, 1948–1950 (1992).
10. Van de Walle, C. G. & Northrup, J. E. *Phys. Rev. Lett.* **70**, 1116–1119 (1993).
11. Van Buuren, T. et al. *Appl. Phys. Lett.* **63**, 2911–2914 (1993).
12. Hybertsen, M. S. *Phys. Rev. Lett.* **72**, 1514–1517 (1994).
13. Baribeau, J.-M., Lockwood, D. J. & Lu, Z. H. *Mater. Res. Soc. Symp. Proc.* Vol. 382 (Materials Research Society, Pittsburgh, in the press).
14. Lu, Z. H. & Yelon, A. *Phys. Rev. B* **41**, 3284–3286 (1990).
15. Lu, Z. H., Baribeau, J.-M. & Jackman, T. E. *Can. J. Phys.* **70**, 799–802 (1992).
16. Lockwood, D. J. et al. *Can. J. Phys.* **70**, 1184–1193 (1992).
17. Tsu, R. *Nature* **364**, 19 (1993).
18. Street, R. A. & Biegelsen, D. K. In *The Physics of Hydrogenated Amorphous Silicon* (eds Joannopoulos, J. D. & Lucovsky, G.) 195–259 (Springer, Berlin, 1984).

ACKNOWLEDGEMENTS. We thank J. P. McCaffrey for TEM measurements, H. J. Labbé for PL measurements, and K. H. Tan for assisting with synchrotron measurements at the US NSF funded Synchrotron Radiation Center. Z.H.L. thanks T. Tiedje for sending his relevant works, and A. Yelon, M. S. Hybertsen and Y. H. Xie for many useful discussions.

Femtosecond molecular dynamics of tautomerization in model base pairs

A. Douhal, S. K. Kim & A. H. Zewail

Arthur Amos Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91125, USA

HYDROGEN bonds commonly lend robustness and directionality to molecular recognition processes and supramolecular structures^{1,2}. In particular, the two or three hydrogen bonds in Watson–Crick base pairs bind the double-stranded DNA helix and determine the complementarity of the pairing. Watson and Crick pointed out³, however, that the possible tautomers of base pairs, in which hydrogen atoms become attached to the donor atom of the hydrogen bond, might disturb the genetic code, as the tautomer is capable of pairing with different partners. But the dynamics of hydrogen bonds in general, and of this tautomerization process in particular, are not well understood. Here we report observations of the femtosecond dynamics of tautomerization in model base pairs (7-azaindole dimers) containing two hydrogen bonds. Because of the femtosecond resolution of proton motions, we are able to examine the cooperativity of formation of the tautomer (in which the protons on each base are shifted sequentially to the other base), and to determine the characteristic timescales of the motions in a solvent-free environment. We find that the first step occurs on a timescale of a few hundred femtoseconds, whereas the second step, to form the full tautomer, is much slower, taking place within several picoseconds; the timescales are changed significantly by replacing hydrogen with deuterium. These results establish the molecular basis of the dynamics and the role of quantum tunnelling.

There are two possible mechanisms of double proton transfer in these model base pairs: a stepwise transfer from the base-pair structure (B-PS) to the tautomer structure (TS) through an intermediate (zwitterionic) structure (IS), or a direct cooperative transfer of B-PS to TS (Fig. 1a). Since the seminal work by Taylor, El Bayoumi and Kasha⁴, experimental evidence of a photoinduced proton-transfer reaction in solution and matrices at 4.2 K (refs 4, 5) as well as in a supersonic-jet expansions⁶, has been reported. In solution, the first picosecond study⁷

showed that the transfer occurs in less than 5 ps. Subpicosecond⁸ studies have recently been reported and showed an excited-state lifetime of 1.4 ps in hexadecane at room temperature. The issue of cooperativity or concertedness of the motion must take into account the role of the solvent and the intramolecular dynamics.

Our experimental system is shown in Fig. 1b. An initial ultra-violet femtosecond (fs) pulse (305–310 nm) excites the base pair (formed in a molecular beam to avoid collisional deactivation and other intermolecular perturbations) to the first electronically excited state, defining a zero of time. A second fs pulse (~620 nm), delayed in time, probes the dynamics of the pair by exciting the bases above their ionization threshold, so that they can then be monitored by time-of-flight (TOF) mass spectrometry. As with other fs studies⁹, the zero of time can be established *in situ* by identifying the ion signal that results when the two pulses coincide. We have also studied the effect of deuterium substitution to examine possible quantum tunnelling effects of the motion.

Figure 2 shows the TOF mass spectra obtained when exciting and probing pulses at no delay ($t=0$); the ionization potential of the monomer is 8.11 eV (ref. 6). The mass peaks at 118 and 236 correspond to the monomer and to the dimer of 7-azaindole, respectively. Figure 2 also gives the corresponding spectra of the deuterated species of NH and CH bonds, assigned on the basis of the stronger acidity of NH relative to CH (ref. 10).

Figure 3 shows the fs transients for the fully undeuterated (NH,NH,CH) pair at two different vibrational energies, E , above the zero-point energy⁶. For the dimer, the decay was fitted to a bi-exponential function, giving decay times of $\tau_1=650$ fs and $\tau_2=3.3$ ps when $E=0$. For $E \approx 1.5$ kcal mol⁻¹, these rates changed significantly giving $\tau_1=200$ fs and $\tau_2=1.6$ ps, indicating the presence of a reaction barrier. (We could not obtain an accurate fit of the full decay when a single exponential behaviour was assumed; Fig. 3.)

For a cooperative, single-step change of the base-pair structure to the tautomer structure, only a single exponential decay is expected, even with a reversible process. In contrast, a two-step process,



will give rise to bi-exponential transients, reflecting the decay (with k_1) of B-PS and the build-up (k_1) and decay (k_2) of IS (note that TS has a nanosecond lifetime, so its decay is not

FIG. 1 *a*, The molecular structures involved in the two-step, or one-step, cooperative double proton transfer dynamics of the base pair (7-azaindole dimer); *b*, schematic diagram of part of the experimental apparatus showing the generation of the femtosecond (λ_1 , pump; λ_2 , probe) pulses (ref. 9, pages 120–134) and the skimmed molecular beam (ref. 9, pages 319–327), together with the detection compartment which uses time-of-flight mass spectrometry. The compressed, amplified pulses were 60 fs in duration and had an energy of ~ 0.5 mJ. The cross-correlation of the pump and probe pulses was measured *in situ* to be ~ 150 fs. The power of the pulses were controlled to study their effect on the transients.

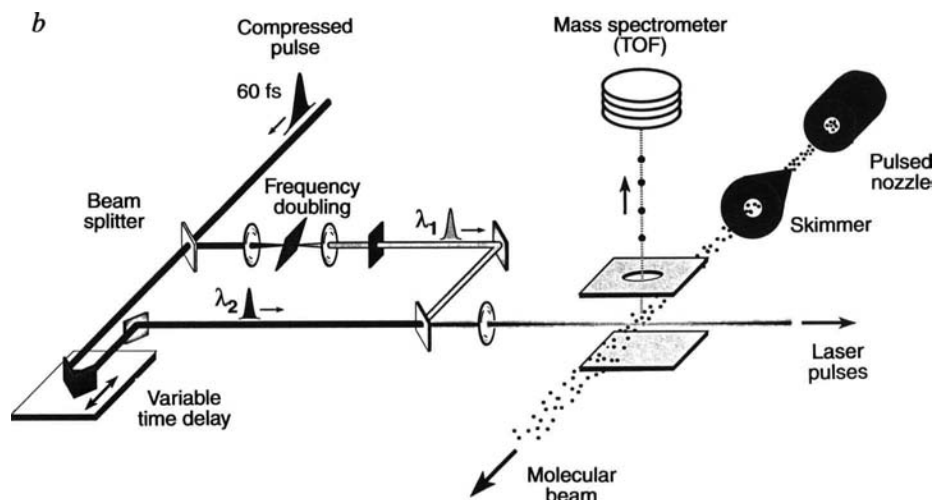
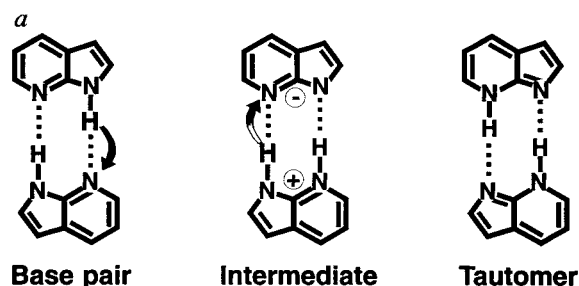
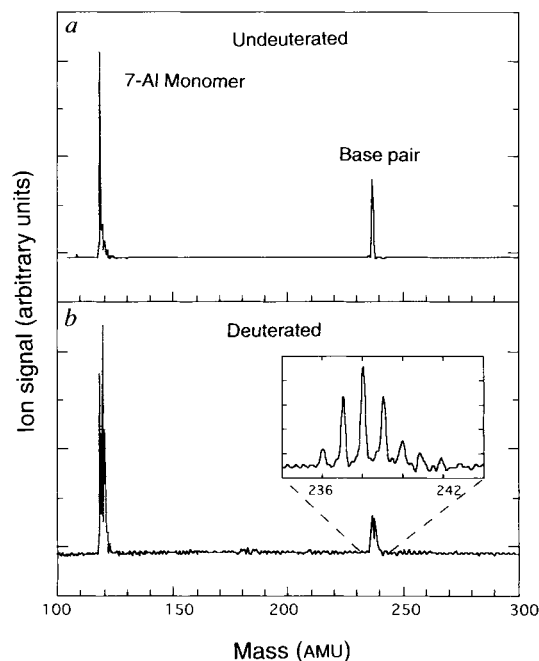


FIG. 2 Time-of-flight mass spectra of the (upper panel) undeuterated and (lower panel) deuterated pair at zero-time delay. The masses at 118 and 236 AMU are due to ions of monomer and base pair, respectively. Mass 236 corresponds to the undeuterated pair, labelled (NH,NH,CH); mass 237 to a mono-deuterated pair, most probably predominantly of type (ND,NH,CH); mass 238, which is the most intense one, is due to the di-deuterated pair, assigned to the (ND,ND,CH) species. The other observed peaks at higher masses (239, 240, 241 and 242) show the existence of 3, 4, 5 and 6 deuterium atoms in the pair indicating the deuteration of both the NH and CH bonds. To form the molecular beam, 7-azaindole (98%, Aldrich) was heated to 373 K with a backing pressure of helium (~ 300 torr). Special care in optimizing the conditions of the beam was made to form the pair; details will be published elsewhere. The skimmer aperture was 2 mm and the nozzle diameter was 0.3 mm. Deuteration was carried out by dissolving the parent compound in a mixture of acetone- D_2O which was later evaporated at 333 K under vacuum.



evident on these timescales). As all of these structures have the same mass, they all contribute to the dimer mass peak, but with different cross-sections for ionization (as in other mass-detection experiments⁹). The fact that the initial tautomerization is on the fs timescale, when the total vibrational energy is zero, indicates that the proton transfer motion is direct and does not involve the entire vibrational phase space of the pair. The implication is that the motion can be described as being 'localized' in the coordinate of N-H...N.

The rate of tautomerization can be modelled by considering the tunnelling motion (rate constant k) of the proton in a double-well potential, with the system either in the N-H or the NH⁺ configuration of the two moieties:

$$k = \nu \exp[-\pi a \sqrt{2mU_0}] \quad (2)$$

where ν is reaction-coordinate frequency (N-H), $a_0 = \hbar a$ is the half-width of the energy barrier and U_0 is its height, and m is the effective mass of the particle. Using the measured k and taking ν for the N-H vibration of 2,800 cm⁻¹ and $a_0 = 0.27$ Å (ref. 5), we obtained $U_0 = 1.2$ kcal mol⁻¹. This model assumes that the distance between the two nitrogens is fixed. When relaxing this condition and averaging over the stretch motion of the N...N centres (two-dimensional model) at $E=0$ we again obtained $U_0 \approx 1.3$ kcal mol⁻¹. For self-consistency, we have repeated these two types of calculations¹¹ and obtained satisfactory agreement for the effect of isotope substitution and also for the excess vibrational energy, as discussed below.

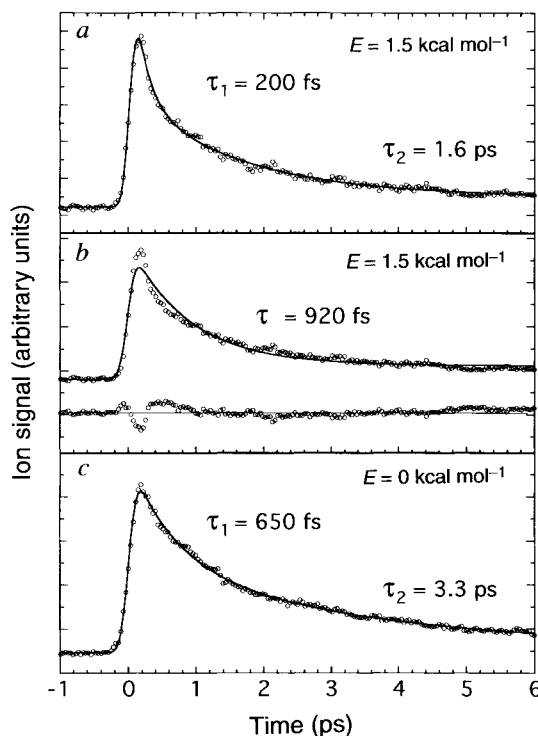


FIG. 3 Femtosecond transients of the base pair (236 AMU) as a function of the time delay. *a*, Excitation with an excess vibrational energy $E \approx 1.5$ kcal mol⁻¹; *b*, a single exponential fit of the same data; and *c*, excitation with $E \approx 0$. A bi-exponential function (and a small constant background) is used to fit both decays in *a* and *c*. The constant takes into account the nanosecond decay, a contribution of $\sim 10\%$, of the total signal. Note the distribution of residuals (lower trace in *b* when a single exponential decay is assumed). The observed transient is not power-dependent, and only one ultraviolet photon excitation is responsible for the observed dynamics.

Löwdin¹² has discussed the theoretical description of the double-well process and its relevance to genetics. One consequence of this process of barrier tunnelling¹³ is the effect on the rates by changing m by deuteration. As shown in Fig. 4, the 360-fs component changes to 3 ps on deuteration. This change is totally consistent with the above calculation when changing $m=1$ (H) to $m=2$ (D). In the two-dimensional model, for example for E corresponding to two quanta of the N...N stretch (0.7 kcal mol⁻¹), we obtained a k^{-1} value of 300 fs for the proton transfer and 5.5 ps for the deuteron transfer, consistent with the experimental trends.

The ps component characterizes the decay of the IS to the final tautomer. From the energy dependence and isotope effect (Figs 3 and 4) we obtain a barrier height using the above procedure; from the rates, a lower limit of 2.6 kcal mol⁻¹ is obtained for this second step. This barrier height is larger than that of the first step, which is reflected in the fact that the characteristic timescale is of the order of picoseconds rather than femtoseconds.

With the equivalence of the two hydrogen bonds in the static structure of the pair it is interesting to consider the nature of the process that leads to the dynamical structures. We propose the following model. Because the timescale of the proton motion is observed to be relatively short, compared to the energy redistribution¹⁴, the 'reaction centre' involves primarily the N-H and N...N intermolecular motions. The timescale of the proton motions, however, is longer than (or comparable to) the changes in the electronic distribution on excitation and the nuclear vibrational motions of the N-H and N...N stretches. This last inequality allows for the asymmetric motion of one of the protons, and, because one moiety is excited, the proton ultimately

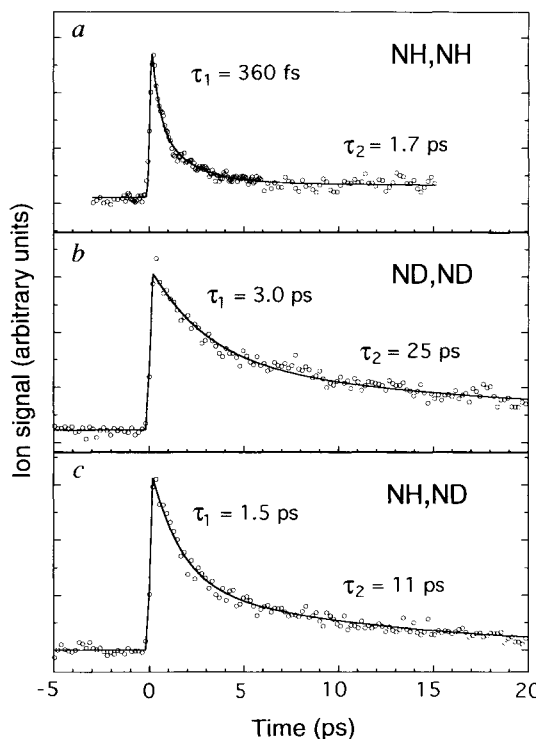


FIG. 4 Femtosecond transients of the ion signal of the base pair (NH,NH,CH) (*a*), (ND,ND,CH) (*b*) and (ND,NH,CH) (*c*) as a function of the time delay with $E \approx 1$ kcal mol⁻¹. To obtain accurate values of the slow component, τ_2 , data from scans with longer time delays (60 ps) were taken and analysed.

transfers, leading to the IS. A consequence of this transfer is stability for the second N–H motion and a higher barrier toward TS formation. The N...N stretch has been assigned⁶ the value 120 cm⁻¹ and the N–H stretch is about 2,800 cm⁻¹, giving 280 and 12 fs, respectively. Therefore, on the timescale of 0.5–10 ps (typical reaction times) the ‘asymmetric reaction coordinate’ for the two particles is established.

To further probe this picture we have examined the transient behaviour of the pair with NH,ND composition. A comparison of the transient behaviour of three species is made in Fig. 4. It is remarkable that the (ND,NH,CH) species gives rates ‘in between’ the values of the (NH,NH,CH) and (ND,ND,CH) species. This behaviour can be understood by considering the breakage of the phase in the vibrational motion of the two N–H bonds. Considering 50% probability for exciting either part of the base pair, we deduce effective time constants of ~1 ps and ~20 ps which describe the experimental trend reasonably well. Interestingly, for the (ND,ND,CD) species, with presumably one deuterium on the carbon rings, the transient behaviour is very similar to that of the (ND,ND,CH) species, again consistent with the locality of the dynamics in the hydrogen-bond regions.

The process of mutation by tautomerization³ is similar to the excited-state process described here. If a ‘misprint’ induced by a tautomer takes place during replication^{15,16}, then an error is recorded. Because reaction-path calculations of DNA base

pairs¹⁷ show similar potential energy characteristics to those discussed here, we anticipate being able to explore the relevance of tautomerization dynamics to mutagenesis. □

Received 7 September; accepted 25 October 1995.

1. Zewail, A. (ed.) *The Chemical Bond: Structure and Dynamics* (Academic, San Diego, 1992).
2. Ball, P. *Designing the Molecular World* (Princeton Univ. Press, 1994).
3. Watson, J. D. H. & Crick, F. H. C. *Nature* **171**, 737–738 (1953).
4. Taylor, C. A., El Bayoumi, M. A. & Kasha, M. *Proc. natn. Acad. Sci. U.S.A.* **63**, 253–260 (1969).
5. Ingham, K. C. & El-Bayoumi, M. A. *J. Am. chem. Soc.* **96**, 1674–1682 (1974).
6. Fuke, K. & Kaya, K. *J. phys. Chem.* **93**, 614–621 (1989).
7. Hetherington, W. M., Micheels, R. H. & Eistenthal, K. B. *Chem. Phys. Lett.* **66**, 230–233 (1979).
8. Share, P., Pereira, M., Sarisky, M., Repinec, S. & Hochstrasser, R. M. *J. Luminescence* **48/49**, 204–208 (1991).
9. Zewail, A. H. *Femtochemistry: Ultrafast Dynamics of the Chemical Bond* (World Scientific, Singapore, 1994).
10. Yu, H.-T., Colucci, W. J., McLaughlin, M. L. & Barkley, M. D. *J. Am. chem. Soc.* **114**, 8449–8454 (1992).
11. Kim, S. K. et al. *J. phys. Chem.* **99**, 7421–7435 (1995).
12. Löwdin, P.-O. *Adv. Quantum Chem.* **2**, 213–361 (1965).
13. Benderskii, V. A. & Goldanskii, V. I. *Int. Rev. phys. Chem.* **11**, 1–70 (1992).
14. Herek, J. L., Pedersen, S., Bañares, L. & Zewail, A. H. *J. chem. Phys.* **97**, 9046–9061 (1992).
15. Cairns, J. J. *molec. Biol.* **6**, 208–213 (1963).
16. Stryer, L. *Biochemistry* 3rd edn (Freeman, New York, 1988).
17. Rein, R. & Harris, F. E. *J. chem. Phys.* **42**, 2177–2180 (1965).

ACKNOWLEDGEMENTS. A. Douhal is on leave from the Universidad de Castilla-La Mancha (Spain), and thanks that university and the California Institute of Technology (USA) for support and for making possible a fruitful stay at Caltech. We thank S. Pedersen for his help with Fig. 1. This work was supported by the US Air Force Office of Scientific Research and the US NSF.

Cell-free synthesis of polyketides by recombinant erythromycin polyketide synthases

Rembert Pieper*, Guanglin Luo†, David E. Cane† & C. Khosla*†

* Department of Chemical Engineering, Stanford University, Stanford, California 94305-5025, USA

† Department of Chemistry, Brown University, Box H, Providence, Rhode Island 02912, USA

MODULAR polyketide synthases (PKSs) are complex multi-enzyme proteins that catalyse the bacterial biosynthesis of many pharmaceutically useful polyketides. The PKSs are organized into a series of modules, each containing the active catalytic sites required for one step in the synthesis process^{1–4}. Here we report a method for cell-free enzymatic synthesis of 6-deoxyerythronolide B (6-dEB), the parent molecule of the antibiotic erythromycin A, using recombinant 6-deoxyerythronolide B synthase (DEBS), a modular PKS with at least 28 distinct active sites. We have also synthesized *in vitro* a triketide lactone by using a truncated mutant of DEBS. The availability of such cell-free synthetic routes will allow direct investigation of the structural and mechanistic basis for the unusual combination of high substrate specificity^{5–10} and tolerance to genetic reprogramming^{2,11–15} found in this enzyme family.

Polyketides are synthesized in a manner analogous to fatty-acid biosynthesis¹⁶ in which carbon chains are built by successive condensations between monomers such as acetate, propionate and butyrate. After each condensation reaction, the resulting β -carbonyl undergoes all, part or none of a reductive cycle comprising β -ketoreduction, dehydration, and enoyl reduction^{17,18}. Thus structural diversity in this family of natural products arises from variations in the choice of monomers, the extent of the β -ketoreduction that occurs during each condensation cycle, the stereochemistry of each chiral carbon centre, and the regio-

chemistry of cyclization reactions that occur after chain synthesis is completed.

The PKS 6-deoxyerythronolide B synthase (DEBS) consists of three multifunctional proteins, DEBS 1, DEBS 2 and DEBS 3, each of which possesses two modules (Fig. 1). The PKS-mediated process of complex polyketide biosynthesis has been explored by radioisotope and stable isotope labelling experiments^{5–10} (see ref. 17 for review), heterologous expression¹⁹, directed mutagenesis^{2,11–15}, and *in vitro* studies on partly active proteins (ref. 20 and references therein). But cell-free enzymatic synthesis of complex polyketides has proved unsuccessful despite more than 30 years of intense efforts^{21–23} (the claim in ref. 24 to have observed cell-free synthesis of 6-dEB has not been followed by any supporting published data), presumably because of the difficulties in isolating fully active forms of these large, poorly expressed multifunctional proteins from naturally occurring producer organisms, and because of the relative lability of intermediates formed during the course of polyketide biosynthesis. The three DEBS proteins have been purified individually from the natural producer organism, *Saccharopolyspora erythraea*²⁵, and studies of the purified enzymes facilitated clarification of their stereospecificity. However, the absence of chain-elongation activity in these preparations prevented detailed investigations of the mechanisms of this enzyme complex. In an attempt to overcome some of these limitations, modular PKS subunits have been expressed in heterologous hosts such as *Escherichia coli* (ref. 26 and references therein) and *Streptomyces coelicolor*¹⁹. Whereas the proteins expressed in *E. coli* are not fully active, heterologous expression of the DEBS PKS in *S. coelicolor* resulted in production of active protein, as demonstrated by the production of 6-dEB *in vivo*. Cell-free enzymatic synthesis of polyketides from PKSs with substantially fewer active sites, such as the 6-methylsalicylate synthase²⁷, chalcone synthase²⁸, tetracenomycin synthase²⁹, and the PKS responsible for the polyketide component of cyclosporin³⁰, have also been reported.

Here we have focused on polyketide synthesis by recombinant DEBS and by an active deletion mutant whose construction has previously been reported^{13,15}. The deletion mutant, designated ‘DEBS 1 + TE’, contains the first two modules from DEBS 1 fused to the thioesterase domain normally found at the carboxy-terminal end of module 6 of DEBS (Fig. 1). Both the complete

† To whom correspondence should be addressed.

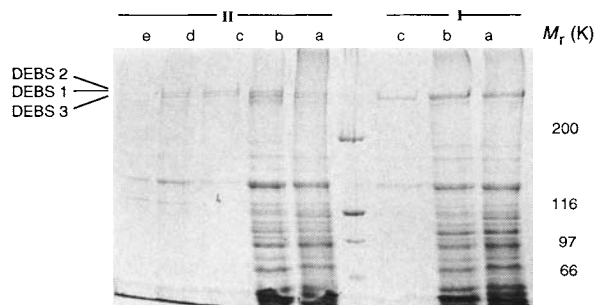
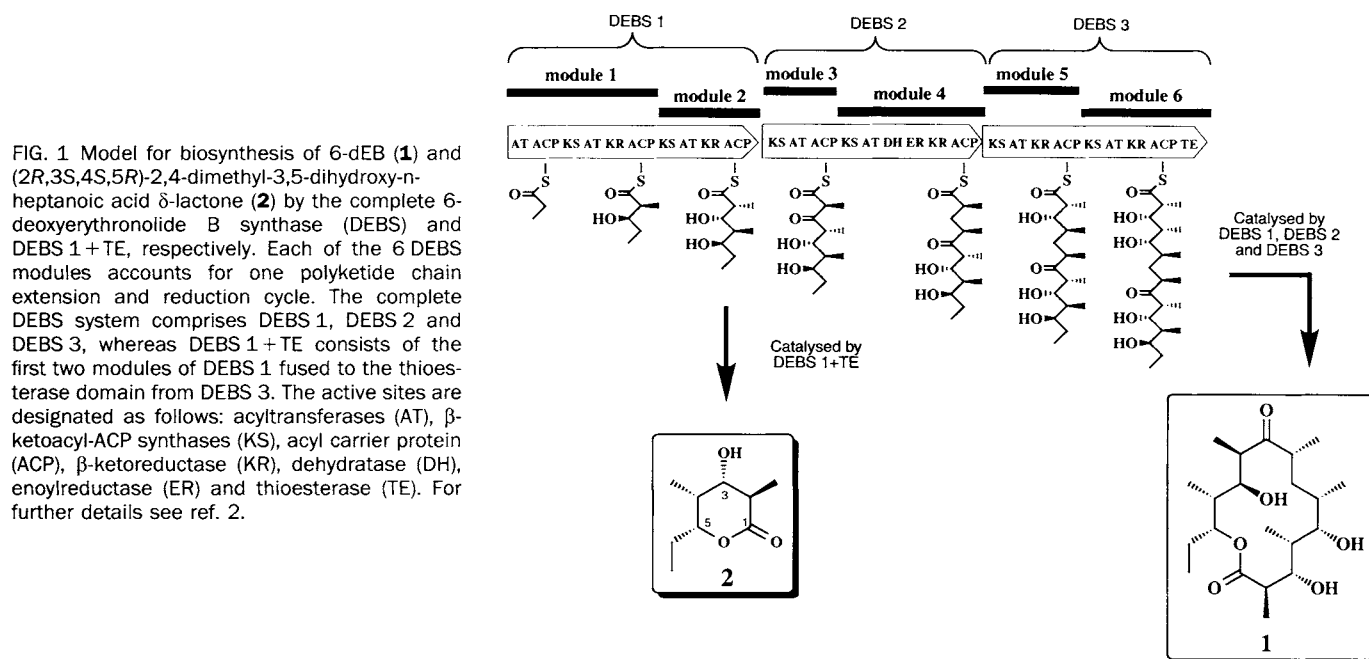


FIG. 2 Partial purification and characterization of the complete DEBS and of DEBS 1+TE from recombinant *S. coelicolor* CH999. Coomassie blue-stained 5% acrylamide gel of protein fractions containing DEBS 1-Thioesterase (TE) (from pCK 12) or the complex of DEBS 1, DEBS 2 and DEBS 3 (from pCK7). The DEBS preparation was performed as follows: *S. coelicolor* CH999/pCK12 or CH999/pCK7 cells were collected after 55 h growth in liquid cultures. Typically, cells of wet weight 8–10 g were disrupted using sonication (five bursts of 30 s). The cell slurry was

ultracentrifuged (2 h at 192,000g) and nucleic acids precipitated with 0.2% polyethyleneimine (step 1) yielding about 200 mg total protein. In a 55%-saturated ammonium sulphate solution all three DEBS proteins were precipitated. The incubation buffer (buffer I) used thereafter contained 150 mM sodium phosphate buffer (pH 7.1), 15% glycerol, 2 mM DTT and 2 mM EDTA. After desalting on Sephadex G25 M (step 2), about 30 mg protein (15–20 mg ml⁻¹) was applied to an Agarose BioGel A size exclusion column (140 ml). Fractions containing DEBS proteins were pooled and concentrated to 1 mg ml⁻¹ on YM 30 ultrafiltration membranes (step 3). After SDS-PAGE DEBS proteins were detected by their high *M_r* values of 330K (DEBS 3), 370K (DEBS 1) and 380K (DEBS 2); these proteins were absent in cell extracts of a variety of control strains. The apparent *M_r* values of the DEBS proteins were also evaluated by gel filtration on a Superose 6 HR 10/30 column (Pharmacia), using thyroglobulin (669K) and apoferritin (443K) as high-*M_r* markers (see text). I, DEBS 1+TE in a after purification step 1, in b after step 2, in c after step 3. II, DEBS 1, 2 and 3 in a after step 1, in b after step 2, in c, d and e after step 3, in order of increasing elution volume. In all lanes containing the DEBS 1 protein, a minor band can be observed that migrates slightly faster than DEBS 3; we believe this protein to be a proteolytic degradation product of DEBS 1, based on its covalent modification by labelled propionyl CoA (data not shown).

DEBS system¹⁹ and DEBS 1+TE^{13,15} have been successfully expressed in *S. coelicolor*, leading to *in vivo* synthesis of 6-dEB (1) and (2R,3S,4S,5R)-2,4-dimethyl-3,5-dihydroxy-n-heptanoic acid δ-lactone (2), respectively. The recombinant strains were also shown to produce 8,8a-deoxyoleandolide and (2R,3S,4S,5R)-2,4-dimethyl-3,5-dihydroxy-n-hexanoic acid δ-lactone, respectively, analogues generated by the incorporation of an acetate primer unit in place of a propionate. The attractiveness of this heterologous expression system as a source for the preparation of an active enzyme *in vitro* was underscored by the observation that labelled substrates and intermediates added exogenously are incorporated into the biosynthetic product with unusual efficiency³¹.

Recombinant DEBS proteins were isolated from cell extracts and partially purified as described in Fig. 2. Size exclusion chromatography of a crude extract containing the three DEBS subunits on Biogel A and Superose 6 (upper size exclusion limit, relative molecular mass 15,000,000 (*M_r*, 15M) and 1.5M, respectively) indicated that DEBS 1 and DEBS 2 associate more tightly with each other than with DEBS 3. Moreover,

DEBS 1 and DEBS 2 (*M_r* 370K and 380K, respectively) elute over a wide range of fractions corresponding to *M_r* between 10M and 1M, indicating that they might form a multimeric complex which partly dissociates during gel filtration. DEBS 3, however, is not present in this extremely large *M_r* range. From size-calibration experiments on a Superose 6 column, DEBS 3 (*M_r* 330K) mostly elutes as a dimer (similar to thyroglobulin, *M_r* 669K). On concentration, the Biogel A column fractions (see Fig. 2) containing DEBS 1 and DEBS 2 alone were found to be inactive *in vitro*. However, when pooled with a concentrated fraction containing DEBS 3 alone (Fig. 2), the reconstituted complex of the three proteins showed comparable activity to the DEBS 1, DEBS 2 and DEBS 3 preparation derived by ammonium sulphate precipitation of the crude cell extract (see below). Our purification results therefore suggest that activity of DEBS requires the formation of a high-molecular-weight oligomeric complex, possibly a trimer of dimers. The formation of homodimers by purified (but not fully active) DEBS subunits has been reported previously (refs 20, 23 and references therein).

Partially purified preparations of the complete DEBS 1, DEBS 2 and DEBS 3 system and of DEBS 1 + TE were incubated with their natural substrates [$1\text{-}^{14}\text{C}$]propionyl-CoA, (2*RS*)-methylmalonyl-CoA, and NADPH. Each protein preparation synthesized a ^{14}C -labelled product with thin-layer chromatography (TLC) retardation factor (R_f) values identical to those of reference samples of either 6-dEB (**1**) or (2*R*,3*S*,4*S*,5*R*)-2,4-dimethyl-3,5-dihydroxy-*n*-heptanoic acid δ -lactone (**2**), respectively, as shown by TLC-autoradiography (Fig. 3). The identities of [^{14}C]-(**1**) and [^{14}C]-**2** were confirmed by dilution of each of the TLC-purified radiolabelled products with authentic unlabelled carrier 6-dEB or triketide lactone and repeated recrystallization of each sample to constant activity. The formation of each lactone product showed an absolute requirement for the relevant protein preparation as well as for methylmalonyl-CoA and NADPH and was inhibited by both *N*-ethylmaleimide and cerulenin, both of which are well-known inhibitors of the condensation reactions of fatty-acid biosynthesis^{32,33}. Based on the observed radiochemical yield of purified product, the formation of 6-dEB catalysed by DEBS 1, DEBS 2 and DEBS 3 was estimated to be 33 pmol per mg total protein. By comparison, the formation of (**2**) by DEBS 1 + TE was 600 pmol per mg total protein.

The specificity of labelling in the triketide lactone product was unambiguously confirmed by preparative scale incubation of [$1\text{-}^{13}\text{C}$]propionyl-CoA with DEBS 1 + TE in the presence of methylmalonyl-CoA and NADPH (Fig. 4*a*). Analysis of the derived product by 100 MHz ^{13}C nuclear magnetic resonance (NMR) showed an enhanced peak at 81.3 p.p.m. corresponding to enrichment at the predicted site, C-5, in **2a** (ref. 12).

DEBS 1 + TE enzyme can also process **3**, the *N*-acetylcysteamine (NAC) thioester of the polyketide chain elongation intermediate, (2*S*,3*R*)-2-methyl-3-hydroxypentanoic acid. Thus incubation of [$1\text{-}^{14}\text{C}$]-**3** (refs 5, 31) with DEBS 1 + TE in the presence of methylmalonyl-CoA and NADPH gave rise to a labelled product which had the chromatographic mobility expected for the triketide lactone (**2**) and which could be recrystallized to constant activity when diluted with unlabelled triketide lactone. The specificity of labelling was unambiguously confirmed by preparative scale incubation of [$2,3\text{-}^{13}\text{C}_2$]-**3** (refs 5, 31) with DEBS 1 + TE, methylmalonyl-CoA and NADPH. The ^{13}C NMR spectrum of the resulting triketide lactone **2b** displayed a pair of enhanced and coupled doublets ($J_{\text{CC}} + 34.5$ Hz) centred at 36.7 and 81.3 p.p.m., corresponding to enrichment at each of the expected sites of labelling, C-4 and C-5, respectively (Fig. 4*b*). This result is consistent with the previously reported incorporation of [$2,3\text{-}^{13}\text{C}_2$]-**3** into the erythromycin macrolide^{23,27} in intact-cell experiments, as well as the results of numerous experiments in which NAC thioesters of advanced intermediates of polyketide chain elongation were shown to be incorporated (albeit inefficiently) into other polyketides⁵⁻¹⁰. Our studies therefore support the speculation that these NAC thioesters are directly loaded onto the appropriate active site in the PKS, and do not require the participation of additional proteins or cofactors. This experiment also directly confirms the ability of the macrolide synthase to recognize exogenously added chain-elongation intermediates and load them correctly on the cognate PKS module for further processing to the natural product.

The use of recombinant protein provided high levels of concentrated protein extracts for these experiments, which may be responsible for the success of our approach. In addition, preliminary data suggest that the high salt concentration of the buffers used enhances the assembly of the multi-enzyme complex, presumably via hydrophobic interactions. The cell-free results reported here, in conjunction with the availability of facile mutagenesis tools, open up new possibilities for studying modular PKS structure and mechanisms. The considerable yield of fully active protein from the recombinant source described here has already begun to allow further detailed analysis of the molecular

recognition features of this remarkable multifunctional catalyst³⁴. It should now be possible not only to make quantitative assessments of substrate specificity and to probe subtle mechanistic details with unparalleled power, but to use NMR and mass spectroscopy to detect previously unobservable intermediates of polyketide chain elongation. Finally, cell-free systems such as the one reported here provide a means for the controlled synthesis of polyketides which might otherwise be inaccessible by *in vivo* engineered biosynthesis. For example, by incorporating butyrate in place of a propionyl or acetyl starter unit, it has been possible to synthesize enzymatically the corresponding analogues of the polyketide lactones that are produced by DEBS and its mutants *in vivo*³⁴. Knowledge of the active site composition and organization of PKS modules might now allow the rational deduction of the structure of the derived polyketide natural product and, indeed, the rational synthesis of entirely

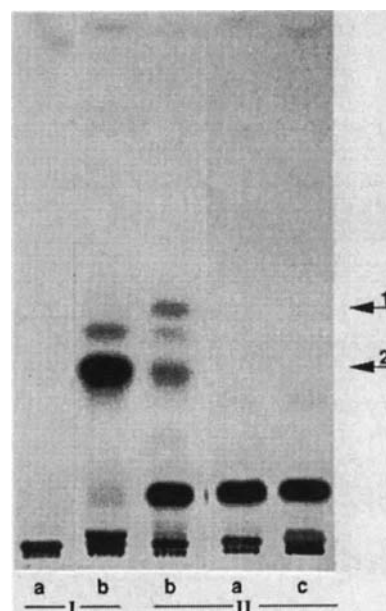


FIG. 3 *In vitro* synthesis of ^{14}C -labelled 6-dEB (**1**) and the triketide lactone (**2**) shown in an autoradiogram. The DEBS 1 + TE (**I**) and DEBS 1, DEBS 2 and DEBS 3 (**II**) preparations described in the Fig. 2 legend (purification step 2) were adjusted to a concentration of 8 mg total protein per ml buffer. Incubations were performed at 28 °C with the [$1\text{-}^{14}\text{C}$]propionyl-CoA (specific activity 50 Ci mol⁻¹, 10 μM), methylmalonyl-CoA (250 μM), and NADPH (500 μM) dissolved in buffer I in a volume of 250 μl for 3 h. Extraction with 2 $\times$ 2 ml ethyl acetate, evaporation of the solvent, and TLC in 60% ethyl acetate/40% hexane followed. The TLC plate was exposed to an X-ray film for 2 days. The autoradiogram shows: **I**, extracts of DEBS 1 + TE, **a**, including [$1\text{-}^{14}\text{C}$]propionyl-CoA and NADPH but excluding methylmalonyl-CoA; and **b**, including all three substrates; **II**, incubations of DEBS 1, DEBS 2 and DEBS 3, **a**, including [$1\text{-}^{14}\text{C}$]propionyl-CoA and NADPH but excluding methylmalonyl-CoA; **b**, including all three substrates; and **c**, preincubation of DEBS 1, DEBS 2 and DEBS 3 with cerulenin (100 μM) for 15 min, followed by addition of all three substrates. Ethyl acetate/hexane (50:50) was used as the solvent system. In lane **Ib** the major labelled component is identical to R_f value (0.30) to authentic triketide **2** (see arrow), whereas in lane **IIb**, the least-polar labelled product is identical in R_f value (0.40) to authentic 6-dEB (**1**) (see arrow). Control experiments were also performed with DEBS 1 + TE in the absence of NADPH, with complete DEBS in the absence of NADPH, and with comparable cell-free preparations from CH999 and from CH999/pSEK38 (a recombinant strain that expresses the actinorhodin PKS gene cluster) in the presence of all substrates and cofactors. In all four controls, neither 6-dEB nor the triketide lactone were detected. The intense, more polar band, evident in all three lanes for **II**, was also present in all of the above null controls, including extracts obtained from CH999 alone. The identities of the enzymatically generated [^{14}C]-**1** and [^{14}C]-**2** were each confirmed by dilution of the respective TLC-purified product with authentic unlabelled carrier and recrystallization to constant activity.

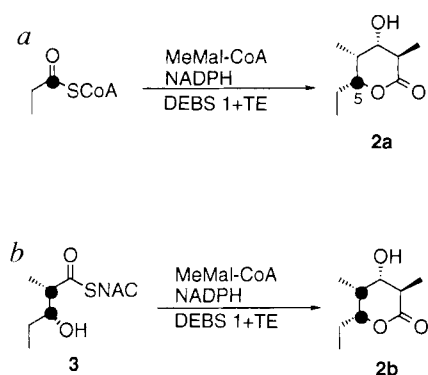


FIG. 4 a, Conversion of [1-¹³C] propionyl CoA to triketide lactone by DEBS 1+TE; b, conversion of [2,3-¹³C₂]-3 to triketide lactone by DEBS 1+TE.

new polyketide structures through the engineering of combinations of PKS modules. Given the structural complexity of most natural or engineered products of modular PKSs 'one-pot' enzymatic synthetic methodologies could be an attractive complement to established chemical synthesis efforts aimed at elucidating the structure-activity relationships of lead molecules with potential human therapeutic, veterinary or agrochemical utility.

Note added in proof: Detection of cell-free activity of DEBS 1+TE has recently been reported elsewhere³⁵. □

Received 25 July; accepted 3 October 1995.

1. Cortes, J., Haydock, S. F., Roberts, G. A., Beville, D. J. & Leadlay, P. F. *Nature* **348**, 176–178 (1990).

2. Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. *Science* **252**, 675–679 (1991).
3. MacNeil, D. J. et al. *Gene* **115**, 119–125 (1992).
4. Schwesche, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 7839–7843 (1995).
5. Cane, D. E. & Yang, C. C. *J. Am. chem. Soc.* **109**, 1255–1257 (1987).
6. Yue, S., Duncan, J. S., Yamamoto, Y. & Hutchinson, C. R. *J. Am. chem. Soc.* **109**, 1253–1255 (1987).
7. Cane, D. E., Prabhakaran, P. C., Tan, W. & Ott, W. R. *Tetrahedron Lett.* **32**, 5457–5460 (1991).
8. Cane, D. E., Lambalot, R. H., Prabhakaran, P. C. & Ott, W. R. *J. Am. chem. Soc.* **115**, 522–566 (1993).
9. Dutton, C. J., Hooper, A. M., Leadlay, P. F. & Staunton, J. *Tetrahedron Lett.* **35**, 327–330 (1994).
10. Cane, D. E. & Luo, G. *J. Am. chem. Soc.* **117**, 6633–6634 (1995).
11. Donadio, S., McAlpine, J. B., Sheldon, P. J., Jackson, M. & Katz, L. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7119–7123 (1993).
12. Kao, C. M., Luo, G., Katz, L., Cane, D. E. & Khosla, C. *J. Am. chem. Soc.* **116**, 11612–11613 (1994).
13. Kao, C. M., Luo, G., Katz, L., Cane, D. E. & Khosla, C. *J. Am. chem. Soc.* **117**, 9105–9106 (1995).
14. Cortes, J. et al. *Science* **268**, 1487–1489 (1995).
15. Brown, M. J. B., Cortes, J., Cutter, A. L., Leadlay, P. F. & Staunton, J. *J. chem. Soc. chem. Commun.* 1517–1518 (1995).
16. Wakil, S. J. *Biochemistry* **28**, 4523–4530 (1989).
17. O'Hagan, D. *The Polyketide Metabolites* (Norwood, Chichester, 1991).
18. Katz, L. & Donadio, S. A. *Rev. Microbiol.* **47**, 875–912 (1993).
19. Kao, C. M., Katz, L. & Khosla, C. *Science* **265**, 509–512 (1994).
20. Marsden, A. F. A. et al. *Science* **263**, 378–380 (1994).
21. Wawszkiewicz, E. J. & Lynen, F. *Biochem. Zbl.* **340**, 213–227 (1964).
22. Corcoran, J. W. in *Antibiotics IV. Biosynthesis* (ed. Corcoran, J. W.) 146–148 (Springer, New York, 1981).
23. Roberts, G. & Leadlay, P. F. *Biochem. Soc. Trans.* **12**, 642–643 (1984).
24. Hunaiti, A. A. & Koliattukudy, P. E. *Antimicrob. Agents Chemother.* **25**, 173–178 (1984).
25. Caffrey, P., Beville, D. J., Staunton, J. & Leadlay, P. F. *FEBS Lett.* **304**, 225–228 (1992).
26. Leadlay, P. F. et al. *Biochem. Soc. Trans.* **21**, 218–222 (1993).
27. Dimroth, P., Walter, H. & Lynen, F. *Eur. J. Biochem.* **13**, 98–110 (1970).
28. Lanz, T., Tropf, S., Marner, F. J., Schröder, J. & Schröder, G. *J. biol. Chem.* **266**, 9971–9976 (1991).
29. Shen, B. & Hutchinson, C. R. *Science* **262**, 1535–1540 (1993).
30. Offenzeller, M. et al. *J. biol. Chem.* **268**, 26127–26134 (1993).
31. Cane, D. E., Luo, G., Khosla, C., Kao, C. M. & Katz, L. *J. Antibiot.* **48**, 647–651 (1995).
32. Plate, C. A., Joshi, V. C. & Wakil, S. J. *J. biol. Chem.* **245**, 2868–2875 (1970).
33. D'Agostino, G. et al. *Biochim. biophys. Acta* **326**, 155–166 (1973).
34. Pieper, R., Luo, G., Cane, D. E. & Khosla, C. *J. Am. chem. Soc.* (in the press).
35. Wiesmann, K. E. H. et al. *Chem. Biol.* **2**, 583–589 (1995).

ACKNOWLEDGEMENTS. This work was supported by the NIH, a National Science Foundation Young Investigator Award (C.K.) and a David and Lucile Packard fellowship for Science and Engineering (C.K.).

Global interdecadal and century-scale climate oscillations during the past five centuries

Michael E. Mann*, Jeffrey Park* & R. S. Bradley†

* Department of Geology and Geophysics, Yale University, PO Box 208109, New Haven, Connecticut 06520-8109, USA

† Department of Geosciences, University of Massachusetts, Amherst, Massachusetts 01003, USA

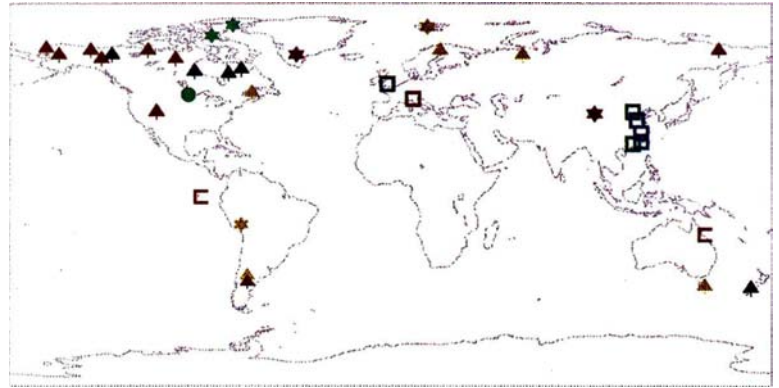
THE recognition of natural modes of climate variability is essential for a better understanding of the factors that govern climate change. Recent models suggest that interdecadal (roughly 15–35-year period)^{1–4} and century-scale (roughly 50–150-year period)^{5–7} climate variability may be intrinsic to the natural climate system. While there is some evidence for the existence of interdecadal^{8,9} and century-scale^{9,10} oscillations in instrumental temperature records, confident detection from these short (100–400-year) records is difficult^{11,12}. Oscillations on the same timescales^{5,13–17} have also been detected in isolated climate-proxy or historical records over longer durations, but the large-scale spatial structure of the variability has not been investigated systematically. Here we report the multivariate analysis of a globally distributed set of temperature proxy records of several centuries duration. The results of our spatio-temporal analysis strengthen evidence for persistent natural interdecadal and century-scale climate oscillations, and reveal both the spatial patterns and temporal histories of these signals.

Evidence for truly global climate signals must demonstrate consistent variations in data from widely separated localities, using records that are long enough to resolve the timescales of interest. To isolate possible climate signals, we have compiled a small (35), but globally distributed, set of high-quality temperature proxy reconstructions (Fig. 1) for the past five centuries. These data include tropical¹⁸ and extratropical¹⁹ ice cores, tropical corals (ref. 14 and J. Lough, personal communication) dendroclimate reconstructions (refs 19–21 and C. Earle, personal communication) and a handful of very long historical records¹⁹. Some seasonal inhomogeneity may be present in these data series because the extratropical records, unlike the tropical records, chiefly reflect summer season variability¹⁹. However, recent analyses²² suggest that large-scale, low-frequency surface temperature variability is largely independent of the seasonality of sampling. Thus we expect that such inhomogeneity will have little impact on signal detection. We employed a frequency-domain singular value decomposition (SVD) method that is designed to isolate and reconstruct quasiperiodic or unstable oscillations that are spatially coherent but which may exhibit phase-lags from site to site^{9,23}.

The sparseness of the data set argues for a cautious interpretation, but the analysis suggests greater potential of our spatio-temporal approach applied to expanded networks of climate-proxy data. We balanced spatial coverage, duration and time resolution (annual versus decadal) by grouping the proxy data into four subsets that were analysed independently: A, 27 shorter records (AD 1730–1969) with annual resolution; B, 21 medium-duration (AD 1615–1969) records with annual resolution; C, 35 shorter (AD 1730–1960) records with decadal resolution; and D, 12 longer (AD 1400–1960) records with decadal resolution.

We compared the proxy network against instrumental data in sampling large-scale climate variability. We performed parallel

FIG. 1 Distribution of proxy temperature reconstructions used in the present study. Squares, historical records; circles, lake varves; stars, ice-core/ice-melt measurements, stemmed triangles, dendroclimate reconstructions; "C" symbols, coral reconstructions. In contrast with other reconstructions, the dendroclimate reconstruction corresponding to the indicated southwestern US 'site' was actually based on the spatial averaging of data over a larger region¹⁹. Records indicated by yellow are available from AD 1400 to 1969 and have annual resolution; records indicated by green are available from AD 1400–1960 and have decadal sampling, red symbols correspond to records available from 1615–1969 with annual resolution, blue symbols records from 1620 to 1960 with decadal sampling, black symbols records from 1730–1969 with annual resolution. Thus, data group A makes use of the 27 yellow, red and black records, data group B the 21 yellow and red records, data group C all 35 records, and data group D the 12 yellow and green records.



analyses of data during the past century based on (1) the global distribution of instrumental temperature data from 1890 to 1989 used by Mann and Park⁹, (2) proxy data set A using the abbreviated interval 1890–1969, and (3) a sparse sub-sampling of the instrumental data during the interval 1890–1969 chosen to mimic the spatial distribution and seasonal sampling of the proxy data. A perfect spatial match between data sets (2) and (3) is not possible, largely due to a paucity of high-latitude instrumental data. The 'SVD spectra' in Fig. 2 show the fraction of data variance in a narrow frequency range that can be explained by a single correlated signal. All three data sets exhibit correlated interdecadal (15–20-year period) variability that we estimate as significant at the 95% confidence level. A roughly decadal (10–12-year) peak corresponding to variability in the high-latitude North Atlantic Ocean and North America during the past century^{9,23} is less consistently observed. Within the secular regime (corresponding to periods greater than the 40 (2) and (3) or 50 (1) years), two distinct modes of variability are significant in each data set, corresponding to a secular warming trend and a weaker variation that resembles one cycle of a century-scale oscillation. The associated spatial patterns of variation (not shown) of these signals in cases (2) and (3) are broadly consistent with their counterparts in (1) shown by Mann and Park⁹. The proxy data network thus seems capable of capturing the large-scale climate processes evident in recent instrumental-based analyses, at least on certain timescales.

The SVD spectra corresponding to the full-length series for data groups A–D (Fig. 3) yield statistically significant peaks on roughly interdecadal (15–35-year) and century (50–150-year) timescales. We isolate a quarter-millennial (~240 year) oscillation in data group D where longer-period variability can be resolved from a secular trend. This identification is tentative, however, as less than three 'cycles' are present, and the signal cannot be independently confirmed from the other data groups. Only groups A and B can properly resolve the interdecadal signals, as the Nyquist frequency for 10-year sampling corresponds to a 20-year period. Higher-frequency signals related to the El Niño/Southern Oscillation (ENSO) that are observed in the A and B analyses²⁴ will be discussed elsewhere. Comparing the spatial patterns of the interdecadal and century-scale signals reveals distinctions that are consistent among the data groups. Peaks of correlated variance fall roughly into interdecadal and century-scale bands. The variation of the location and magnitude of the peaks within these bands suggests quasiperiodic signals with amplitudes and frequencies that drift over time.

The time-evolving amplitude and frequency of quasiperiodic signals can be examined with an 'evolutionary' analysis (Fig. 4), in which the SVD analysis is applied in a moving window through the data series. The short duration, combined with decadal resolution, precludes a meaningful evolutionary analysis for data group

C. Evolutionary analyses of the other data groups demonstrate that coherent interdecadal oscillations, centred near 20–25-year periodicity, were weakly evident before 1800. The oscillations subsequently strengthen in significance and drift to roughly 16–18-year period in the final window (1869–1969), which agrees with the dominant timescale of oscillations observed in instrumental surface-temperature data⁹ during the past century. Time windows that resolve century-scale variations (200-year width) can be employed for data group D. Before 1650, a coherent signal with roughly 50-year period appears intermittently. After 1650 this oscillation strengthens in its significance and drifts to a 60–70-year periodicity. In the nineteenth century, these 'century-scale' oscillations seem to drift to slightly longer period, becoming indistinguishable from secular variability in a 200-year moving window. The improved frequency resolution of data group D allows for a clearer separation of the interdecadal and century-scale variability than is possible with groups A and B.

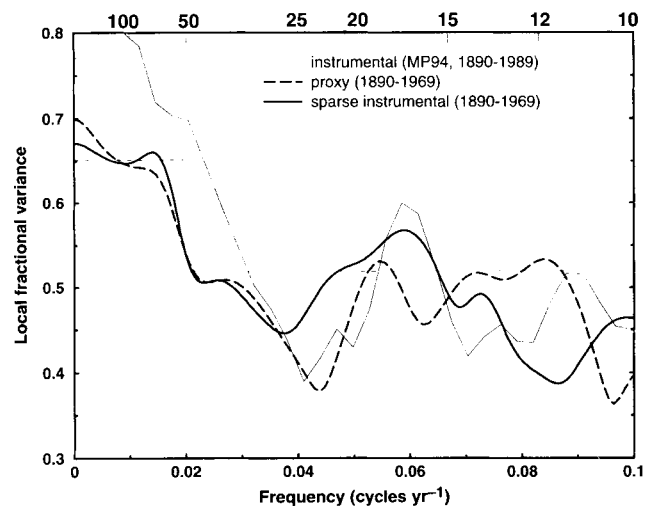


FIG. 2 SVD spectra (local fractional variance explained by principal mode in the SVD within a narrow band about a specified carrier frequency as a function of frequency—see Fig. 3) for (1) instrumental data of Mann and Park⁹ (thin solid line), (2) proxy data set A (dashed line), (3) sparse distribution of instrumental data as described in the text (thick solid line). The local fractional variance scale corresponds to (1), with the scales adjusted for (2) (–0.03), (3) (–0.07) to align the position of the 95% confidence level owing to varying effective spatial degrees of freedom in the different analyses. Confidence limits are higher within the secular band (corresponding to variability longer than half the record lengths) owing to decreased spectral degrees of freedom near zero frequency.

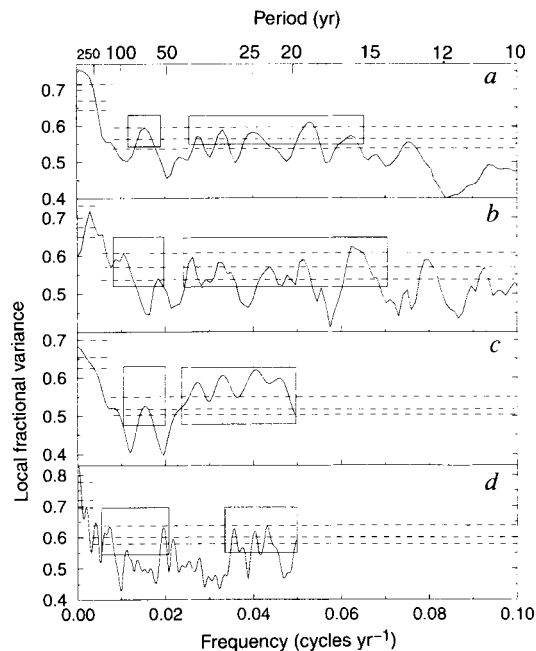


FIG. 3 SVD spectra for each of the data group. A–D. (see Fig. 1 legend) The spectra are plotted as in Fig. 2 but with the 90%, 95%, and 99% confidence limits (dashed lines) each shown. Boxes group spectral peaks into ~15–35-year interdecadal, and ~50–150-year century-scale, bands.

METHODS. We calculate the tapered Fourier transform of a set of M time series of uniform length of N years, using multitaper spectral analysis to provide a small number (K) of independent spectral estimates for each time series $x_{(m)}$ in a narrow frequency band centred at frequency f ,

$$y_k^{(m)}(f) = \sum_{n=1}^N w_n^{(k)} x_{(m)} e^{i2\pi f n \Delta t}$$

where $\Delta t=1$ month is the sampling interval and $\{w_n^{(k)}\}_{n=1}^N$ is the k th member in an orthogonal sequence of leakage-resistant data tapers. We construct the $M \times K$ matrix

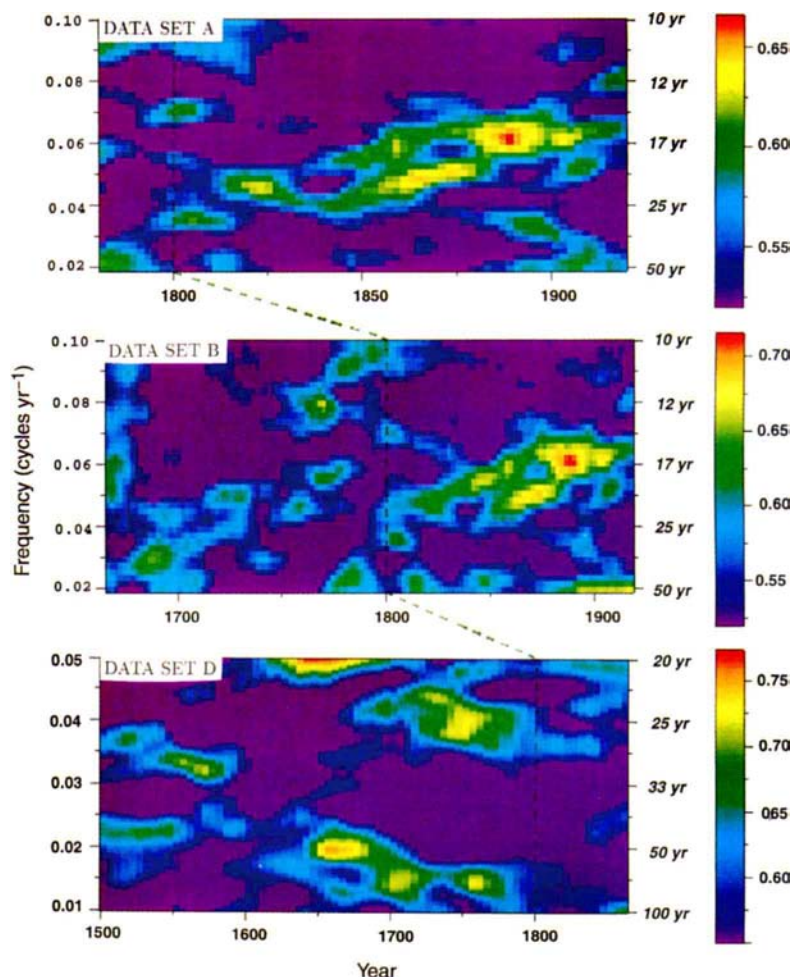
$$\mathbf{A}(f) = [\omega_m y_k^{(m)}(f)]$$

where $m=1, \dots, M$, $k=1, \dots, K$, each row calculated from a different grid-point time series (normalized by its standard deviation). We use $K=3$ as a trade-off between spectral resolution and degrees of freedom. The relative weights for different series (ω_m) are chosen so that a large number of nearby sites do not combine to have disproportional weight in the analysis. We perform the complex singular value decomposition,

$$\mathbf{A}(f) = \sum_{k=1}^K S_k(f) \mathbf{u}_k(f) \otimes \mathbf{v}_k^*(f)$$

into K orthonormal left ($\mathbf{u}_k^*$) and right ($\mathbf{v}_k$) eigenvectors, which can be used to recover spatial and temporal patterns, respectively. The singular value $S_k(f)$ scales the amplitude of the k th mode. The spectrum of 'local' fractional variance explained by the principal mode is described by $S_1^2(f)/\sum_{k=1}^K S_k^2(f)$, with narrow peaks in the spectrum indicating potential quasiperiodic signals. Confidence limits for peak significance are obtained by resampling methods⁹.

FIG. 4 Evolutive SVD spectra. A common year (AD 1800) is indicated by a thick dashed line to allow visual alignment of the time axes for the three plots shown. For each case, the local fractional variance spectrum is indicated by a colour scale shown to the right of the plot. The colour convention is chosen so that any features clearly visible above the background are significant well above the 95% level. Window timelengths τ_w are chosen to be long enough to allow high-frequency resolution, while being short enough to provide insight into the evolving character of signals. The centre of the moving windows defines its time coordinate, so information for the first and last $\tau_w/2$ years are lost. Periods longer than $\tau_w/2$ cannot be distinguished from secular variability in the evolutive analyses, and are not shown. Top panel, Evolutive SVD spectrum for data group A using a 100-year moving window evaluated in 2-year steps. Century-scale variability cannot adequately be distinguished from significant secular (>50-year timescale) variability for this data group. Significance levels for the local fractional variance spectrum are given by 90% (0.55), 95% (0.565) and 99% (0.60). Middle panel, Evolutive SVD spectrum for data group B using a 100-year moving window evaluated in 2-year steps. The significance levels are the same as for data group A. Bottom panel, Evolutive SVD spectrum for data group D using a 200-year moving window evaluated in 5-year steps. For this data subset, century-scale oscillations are largely resolved (although variability on timescales longer than 100 years is indistinguishable from the secular band). Significance levels are given by 90% (0.58), 95% (0.60) and 99% (0.64). Note the different frequency/period scale in this case.



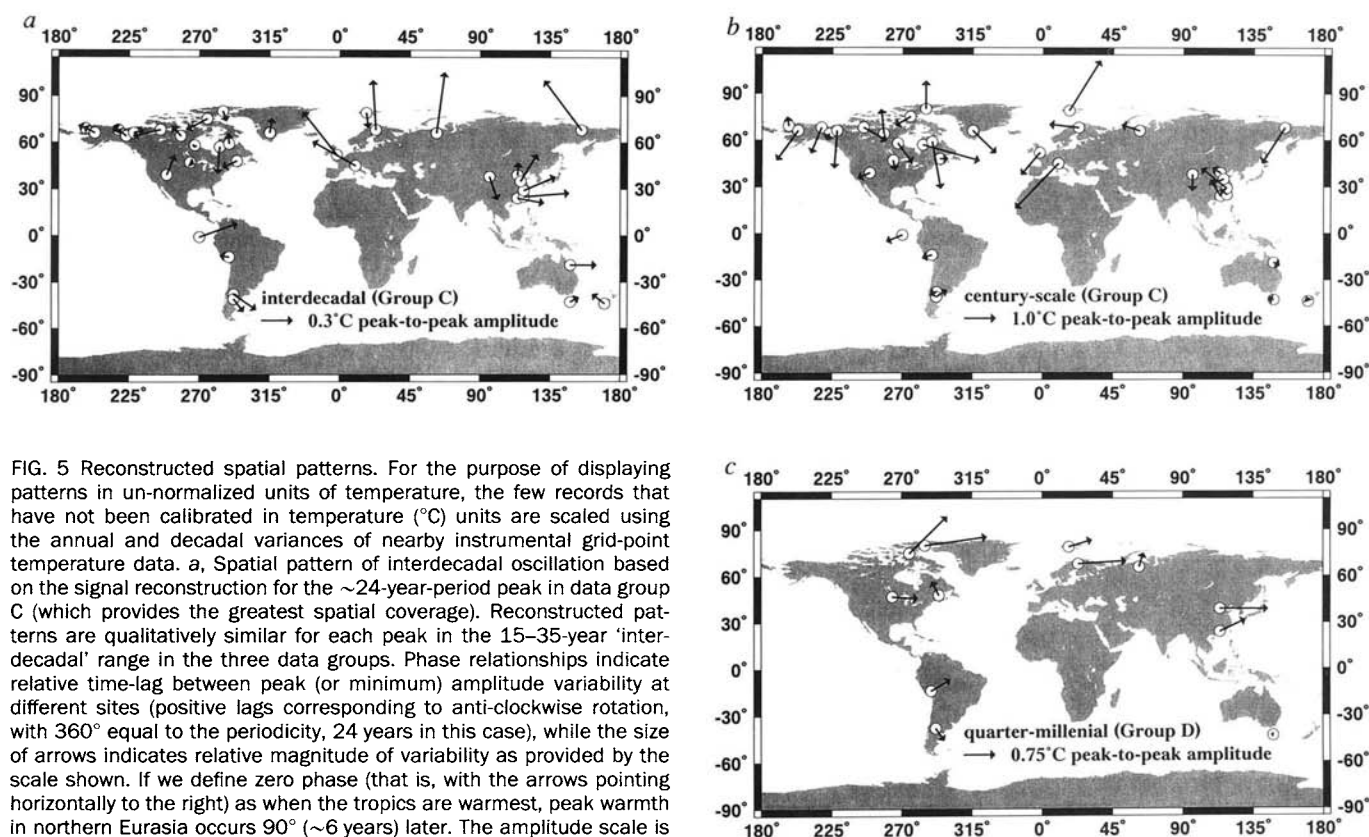


FIG. 5 Reconstructed spatial patterns. For the purpose of displaying patterns in un-normalized units of temperature, the few records that have not been calibrated in temperature ($^{\circ}\text{C}$) units are scaled using the annual and decadal variances of nearby instrumental grid-point temperature data. *a*, Spatial pattern of interdecadal oscillation based on the signal reconstruction for the ~ 24 -year-period peak in data group C (which provides the greatest spatial coverage). Reconstructed patterns are qualitatively similar for each peak in the 15–35-year ‘interdecadal’ range in the three data groups. Phase relationships indicate relative time-lag between peak (or minimum) amplitude variability at different sites (positive lags corresponding to anti-clockwise rotation, with 360° equal to the periodicity, 24 years in this case), while the size of arrows indicates relative magnitude of variability as provided by the scale shown. If we define zero phase (that is, with the arrows pointing horizontally to the right) as when the tropics are warmest, peak warmth in northern Eurasia occurs 90° (~ 6 years) later. The amplitude scale is set so that the largest arrow corresponds to the regional maximum amplitude of the oscillation of $\sim 0.6^{\circ}\text{C}$ peak-to-peak. *b*, Spatial pattern for century-scale signal oscillation reconstructed for the 65-year-period variance peak of data group C. Plot conventions are analogous to those in *a*, with 360° phase difference associated with a 65-year period. The pattern maximum is $\sim 2^{\circ}\text{C}$ peak-to-peak oscillation (for example, central

Europe). *c*, Spatial pattern for the quarter-millennial oscillation. Variations are largely in phase, spread out (with a few exceptions) over ~ 90 – 120° in phase (that is, 60–80 years). The pattern maximum is a $\sim 1.5^{\circ}\text{C}$ peak-to-peak oscillation.

Figure 5 shows reconstructions of the spatial patterns of variation for the interdecadal (panel *a*), century-scale (*b*) and quarter-millennial (*c*) signals. In the spatial reconstructions, sources of systematic bias in the individual temperature reconstructions may lead to unreliable phase and amplitude at specific sites. Calibration of proxy data at longer periods, and various corrections that are made to proxy records (for example, subtraction of individual long-term growth trends for dendroclimate records) are potential sources of such bias. Thus, it is the regional trends evident in these patterns that are most meaningful.

The interdecadal signal (Fig. 5*a*) exhibits variability in the tropics and subtropics that is largely in-phase. Mid-latitude variations are of similar magnitude, but phase relationships are more variable, consistent with the signature of extratropical teleconnection patterns. Such patterns can be resolved by our data network only in part. Nevertheless, the alternating pattern of phase from Greenland, to eastern and then western mid-latitude North America (with a nodal point in central North America) is consistent with the alternating warm and cold advection of the three-lobed ‘Pacific North American’ pressure anomaly pattern. In-phase tropical/extratropical temperature variability coupled with such an extratropical pattern has been noted as a signature of the global-scale ENSO phenomenon in instrumental records^{9,25}. Studies of both instrumental climate data^{9,23,26} and isolated long-term proxy records²⁷ suggest that interdecadal climate variability is closely associated with the Pacific North American pattern. A connection has also been found with low-frequency changes in ENSO²⁶. Although the dynamics of such a linkage are still unclear, our study indicates a possible long-term connection between interdecadal climate oscillations and

low-frequency changes in ENSO. Recent model simulations indicate that the origin of this variability may lie in extratropical ocean–atmosphere interaction³.

The pattern of the century-scale mode (Fig. 5*b*) exhibits high-amplitude variability largely confined to the North Atlantic and Arctic, out-of-phase with weaker variability in the Pacific basin. These features resemble the pattern of a single ‘oscillation’ detected in a century of gridded surface temperature data^{9,10}. Apart from the Svalbard site in the boreal Atlantic, there is a tendency towards opposing, though not opposite, phase anomalies (that is, $\sim 45^{\circ}$ – 135° phase difference) between the eastern and western margins of the North Atlantic. Such a phase pattern suggests a combination of in-phase and out-of-phase components in the two regions. An out-of-phase component could arise from differential temperature advection on either side of an alternating centre of low and high pressure over the North Atlantic. Changes in northward oceanic heat transport would generate an in-phase component of basinwide warming and cooling. The combination of these effects is consistent with coupled ocean–atmosphere model simulations⁷, in which century-scale (40–60-year) oscillations in sea-level pressure over the North Atlantic coincide with oscillations in the thermohaline circulation.

The quarter-millennial oscillation has a spatial pattern that is largely in-phase globally, although the phase-lags in the reconstruction correspond to time-lags as large as 100 years between warm or cold periods at different sites. The time reconstruction (not shown) suggests that this oscillation describes a global-scale transition from warmer conditions before 1500 to the colder conditions of the sixteenth to eighteenth centuries^{19,28}.

We find some empirical evidence for persistent, if variable, low-frequency global-scale climate oscillations on interdecadal

and century timescales. The spatial reconstructions of these signals provide material for hypotheses about the possible underlying processes, for example, a connection between interdecadal climate variations and low-frequency ENSO variability, and a connection between persistent century-scale oscillations and low-frequency variability in the thermohaline circulation. Although such hypotheses are speculative given the limited spatial distribution of the present dataset, they can be tested more rigorously with similar analyses of the expanding network of high-quality climate-proxy reconstructions. □

Received 23 February; accepted 25 September 1995.

1. Barnett, T. P., Del Genio, A. D. & Ruedy, R. A. *J. geophys. Res.* **97**, 7341–7354 (1992).
2. Von Storch, J. S. *Tellus* **46**, 419–432 (1994).
3. Latif, M. & Barnett, T. P. *Science* **266**, 634–637 (1994).
4. Chen, F. & Ghil, M. *J. phys. Oceanogr.* (in the press).
5. Stocker, T. F. & Mysak, L. *Clim. Change* **20**, 227–250 (1992).
6. Mysak, L. A., Stocker, T. F. & Huang, F. *Clim. Dyn.* **8**, 103–116 (1993).
7. Delworth, T., Manabe, S. & Stouffer, R. J. *J. Clim.* **6**, 1993–2011 (1993).
8. Ghil, M. & Vautard, R. *Nature* **350**, 324–327 (1991).
9. Mann, M. E. & Park, J. *J. geophys. Res.* **99**, 25819–25833 (1994).
10. Schlesinger, M. E. & Ramankutty, N. *Nature* **367**, 723–726 (1994).
11. Elsner, J. B. & Tsonis, A. A. *Nature* **372**, 507–508 (1994).
12. Schlesinger, M. E. & Ramankutty, N. *Nature* **372**, 508–509 (1994).
13. Diaz, H. & Pulwarty, R. S. *Clim. Change* **26**, 317–342 (1994).
14. Dunbar, R. B., Wellington, G. M., Colgan, M. W. & Glynn, P. W. *Paleoceanography* **9**, 291–315 (1994).
15. Linsley, B. K., Dunbar, R. B., Wellington, G. M. & Mucciarone, D. A. *J. geophys. Res.* **99**, 9977–9994 (1994).
16. Plaut, G., Ghil, M. & Vautard, R. *Science* **268**, 710–713 (1995).
17. Thomson, D. J. *Phil. Trans. R. Soc. Lond. A* **330**, 601–616 (1990).
18. Thompson, L. in *Climate Since A.D. 1500* (eds Bradley, R. S. & Jones, P. D.) 517–548 (Routledge & Kegan Paul, Boston, 1992).
19. Bradley, R. S. & Jones, P. D. *Holocene* **3**, 367–376 (1993).
20. Jones, J. D. & Briffa, K. R. *Holocene* **1**, 165–179 (1992).
21. Jacoby, G. C. & D'arrigo, R. *Clim. Change* **14**, 39–59 (1989).
22. Briffa, K. R. & Jones, P. D. *Holocene* **3**, 82–92 (1993).
23. Mann, M. E., Lall, U. & Saltzman, B. *Geophys. Res. Lett.* **22**, 937–940 (1995).
24. Bradley, R. S., Mann, M. E. & Park, J. *Eos (suppl.)* **75**, 383 (1994).
25. Halpert, M. S. & Ropelewski, C. F. *J. Clim.* **5**, 577–593 (1992).
26. Trenberth, K. *Bull. Am. met. Soc.* **71**, 988–993 (1990).
27. Slowey, N. C. & Crowley, T. J. *Geophys. Res. Lett.* **22**, 2345–2348 (1995).
28. Hughes, M. & Diaz, H. F. *Clim. Change* **26**, 109–129 (1995).

ACKNOWLEDGEMENTS. We thank G. Jacoby and L. Thompson for providing several data series used in this study, and B. Saltzman for comments on the manuscript. This work was supported by the US NSF (M.E.M., J.P. and R.B.) and the US Department of Energy (R.B.).

The 'Daly gap' as a magmatic catastrophe

C. C. Bonnefoi*†, A. Provost* & F. Albarède†

* Université de Clermont-Ferrand, Département des Sciences de la Terre, 63038 Clermont-Ferrand Cedex, France

† Ecole Normale Supérieure de Lyon, 69364 Lyon Cedex 7, France

IGNEOUS rocks very commonly show a strongly bimodal distribution of compositions, one mode corresponding to basalt and the other to felsic magmas^{1–3}. As fractional crystallization of basaltic parents produces a continuum of compositions, the paucity of rocks of intermediate composition—commonly called the Daly gap—has puzzled petrologists since the time of Daly. Gravitational or viscous trapping^{4,5}, large crystal loads restraining convection^{6,7}, and re-melting of deep volcanic layers² are among the processes that have been offered as physically meaningful explanations of magmatic gaps. Here we propose an alternative interpretation, transposed from chemical reactor control theory⁸: at large undercooling, thermal feedback in a continuously fed and differentiating magma reservoir promotes the existence of competing thermochemical steady states. Small variations in magma residence time and cooling rate induce a large thermal and chemical swing (magmatic bifurcation or catastrophe), which interrupts the liquid line of descent, leading to bimodal erupted products.

We consider a differentiating magma reservoir continuously erupting, crystallizing, and replenished (Fig. 1). The temperature

and composition of the liquid are assumed to be homogeneous within the reservoir, which implies vigorous convection at rather high Rayleigh numbers such as in vertical dykes. Input magmas are crystal-free. The output at first consists of erupted lavas, which initially are assumed to be aphyric (crystal-free), although this assumption is in no way critical. In addition, a mass of cumulate mush fills up the bottom of the reservoir while continuing replenishment by fresh magma is accommodated by the wall rocks being moved apart. The reservoir is defined as the liquid part of the magmatic system. Crystallization changes liquid composition and releases latent heat while the reservoir loses heat by conduction or hydrothermal convection to its surroundings. In our analysis, the following symbols are used: t for time, T for temperature, M for the mass of a phase, Q for magma input/output rate, C_p for heat capacity, L for latent heat, and Φ for the rate of heat loss. Liquid and solid are referred to by the subscripts _{liq} and _{sol}, respectively, input and output by the subscripts _{in} and _{out}.

Mass balance between input, output, and crystal accumulation on the floor and walls of the reservoir requires

$$\frac{dM_{liq}}{dt} = Q_{in} - Q_{out} - \dot{M}_{sol} \quad (1)$$

where the overdot refers to the rate of change. Similarly, the heat-conservation condition is given by

$$\frac{d[M_{liq}(C_p T_{liq} + L)]}{dt} = Q_{in}(C_p T_{in} + L) - Q_{out}(C_p T_{liq} + L) - \dot{M}_{sol} C_p T_{sol} - \Phi \quad (2)$$

Temperature is assumed to be identical in the liquid and in the last increment of cumulate. For constant T_{in} , these equations are reformulated accordingly with respect to a single state-variable, the dimensionless cooling $v = 1 - T_{liq}/T_{in}$ as

$$\frac{dv}{dt} = -\frac{v}{\tau} + \frac{\eta + 1}{\theta} - \eta \dot{f} \quad (3)$$

where $\dot{f} = \dot{M}_{sol}/M_{liq}$ is the fractional rate of solidification. The system has three control parameters: $\tau = M_{liq}/Q_{in}$ the magma residence time in the reservoir, $\theta = M_{liq}(L + C_p T_{in})/\Phi$ the Newtonian cooling time, and $\eta = L/(C_p T_{in})$ the reduced latent heat.

If all control parameters stay constant, the system approaches a steady state with a relaxation time on the order of τ . It is not thus implied that natural systems are time-independent, but this assumption provides a convenient illustration of how a magmatic catastrophe may develop over periods longer than τ . To understand the source of instabilities, we investigate how the thermal regime of the reservoir depends on the rate of replenishment. Equation (3) at thermal steady-state is rewritten as a balance between latent heat release and advection-conduction effects

$$\eta \dot{f} = \frac{\eta + 1}{\theta} - \frac{v}{\tau} \quad (4)$$

Both linear crystal growth rate G and nucleation rate I cancel at the liquidus and below the glass transition. Upon cooling, both G and I therefore go through a maximum at an intermediate temperature⁹. Deferring justification to later in the Letter, we assume that the rate of solidification $\dot{f}$ is also a bell-shaped function of temperature. The righthand side of equation (4) is the equation of a straight line Δ with slope $-1/\tau$, vertical intercept $(\eta + 1)/\theta = \Phi/(M_{liq} C_p T_{in})$, and horizontal intercept $(\eta + 1)\tau/\theta$ (Fig. 2a). Any intersection between the curve $\eta \dot{f}(v)$ and Δ is a solution of equation (4). For very high cooling rates, $(\theta = \theta_1$ in Fig. 2a), there is only one such intersection regardless of the magma residence time τ . For slower cooling $(\theta = \theta_2)$, increasing the magma residence time shifts the system from one to three intersections, then from three to one. The system bifurcates when the straight line Δ becomes tangent to the curve $\eta \dot{f}(v)$. In the case of three intersections A, B and C, the intermediate steady

state B must be unstable, that is, any thermal perturbation however small drives the system away from its initial condition. This is shown by finding the sign of the derivative with respect to v of equation (3) at steady-state and constant η and θ , that is, the sign of

$$\frac{\partial \dot{v}}{\partial v} = -\frac{1}{\tau} - \eta \frac{\partial \dot{f}}{\partial v} = -\frac{v}{\tau^2} \left(\frac{\partial v}{\partial \tau} \right)^{-1} \quad (5)$$

where $\partial v / \partial \tau$ refers to equation (4) read as a relation between v and τ . When $\partial \dot{v} / \partial v$ is negative, any thermal fluctuation decays. When it is positive, any fluctuation grows extremely fast. The middle expression represents the difference between the slope of Δ and that of the tangent at the point of intersection. Points A and C therefore represent a stable steady state, whereas B is unstable. Equivalently, the steady state is stable for $\partial v / \partial \tau \geq 0$ and unstable otherwise. For even slower cooling rates ($\theta = \theta_3$), increasing τ leads to one bifurcation from one to three steady states but not to the other (from three to one). Clearly, the two conditions of instability are efficient heat loss to the surroundings and strong undercooling, when latent heat liberated by solidification acts to enhance solidification rate.

We now describe how a steady-state reservoir may become unstable when the magma replenishment rate Q_{in} and therefore the residence time τ , are changed. We suppose that the cooling time θ lies in the critical interval such as for θ_2 in Fig. 2a. Alternative assumptions, such as a coupled variation of cooling and replenishment rates, would change the analysis somewhat but would not alter the conclusions about the origin of instabilities. When magma input is vigorous (short τ), the slope of Δ is steep and the magma is barely differentiated. When replenishment wanes slowly enough for the reservoir to remain at near

steady-state, temperature decreases along the branch Oa in the (τ, v) space of Fig. 2b, but the thermal regime remains stable as $\partial v / \partial \tau > 0$. At α , magma input cannot be turned further down unless the system abruptly chills down to point α' on the low-temperature branch $\beta\beta$, where solidification may be progressively completed. The temperature discontinuity at the second bifurcation $\alpha\alpha'$ is a true catastrophe in the sense that an infinitely small variation of the control parameter τ triggers a finite change of the system state. Reversing the process by increasing magma input, the reservoir starts warming up along $\beta\beta$ while $\partial v / \partial \tau \geq 0$. At β , it thaws abruptly upon a sharp temperature increase that brings the conditions to those of point β' where further progressive heating can take place. At no point has the system seen the unstable branch $\alpha\beta$. The thermal path is in part not reversible, a feature that amounts to thermal hysteresis. Regardless of the exact compositional relations between coexisting solid and liquid phases, a large temperature drop and massive crystallization must be accompanied by a steep change of the residual magmas towards more felsic compositions. This is true even though the freezing rate is significantly slower at large undercoolings. Some chemical hysteresis of lavas and cumulates must therefore result from the thermal hysteresis.

The dependence of $\dot{f}$ on temperature is crucial to our analysis. The quantity $\dot{f}$ is a non-decreasing function of the linear growth rate and nucleation rate which are both bell-shaped with a sub-solidus minimum^{9,11}. It could in principle be retrieved from temperature-time transformation (TTT) curves which unfortunately have been determined for lunar basalt compositions only¹². At small undercoolings, $\dot{f}$ increases when temperature decreases because undercooling is the dominant driving force. At large undercooling, $\dot{f}$ decreases with temperature because of rising viscosity which restrains ion migration. Dependence of both quantities on magma composition and the existence of a crystallization interval further complicate the problem.

A necessary condition for efficiently cooled magma conduits and reservoirs to thermally bifurcate is that their newtonian cooling time increases beyond a critical value θ_c and that their temperature drops below a critical temperature T_c determined by the intersections of the tangent to the curve $\eta \dot{f}(v)$ at its inflection point with both vertical and horizontal axes. Taking $\eta \approx 0.5$, θ_c is approached for $1/\dot{f} \approx \theta/3$, that is, when the characteristic time for crystallization is somewhat shorter than the newtonian cooling time, which is probably a rather common situation. Although experimental data are few, maximum crystal growth and nucleation rate is only reached in magmatic systems for undercooling in the range of 50–200 K (ref. 9), that is, for a value of $v = v_{max} \approx 0.10 \pm 0.05$. A conservative range of $v_c = 2v_{max} \approx 0.2 \pm 0.1$ is therefore quite acceptable for most magmatic compositions and agrees with the estimate used in ref. 13. Competing steady states may therefore appear for $\tau/\theta > v_c$. Such a situation is probably rather unusual for the basaltic stage of large volcanoes such as those of Hawaii, Réunion or Mt Etna, for which residence times commonly seem to be shorter than a few tens of years^{14,16}. In contrast, early or waning stages of these volcanoes usually have small production of differentiated magmas of alkalic affinity such as hawaiites, mugearites and phonolites¹⁷. These volcanic stages should therefore be associated with long magma residence times, and strong conductive plus hydrothermal cooling, and be prone to thermal instabilities and chemical gaps. The Daly gap has indeed been best defined in this very specific igneous environment^{1,3}. The strontium-isotope composition of the tephrites erupted historically by Vesuvius becomes less radiogenic with time, which was shown^{15,18} to indicate a magma residence time of about 1,000 years and a reservoir volume of less than one cubic kilometre. Admitting for the sake of illustration that this reservoir is a 100-metre-thick dyke, standard equations of heat transfer¹⁹ can be used to infer that such a well stirred reservoir should freeze in a few thousand years by dry conduction. The Vesuvius reservoir seems to have a rather high τ/θ ratio and therefore offers all the potential for a magmatic

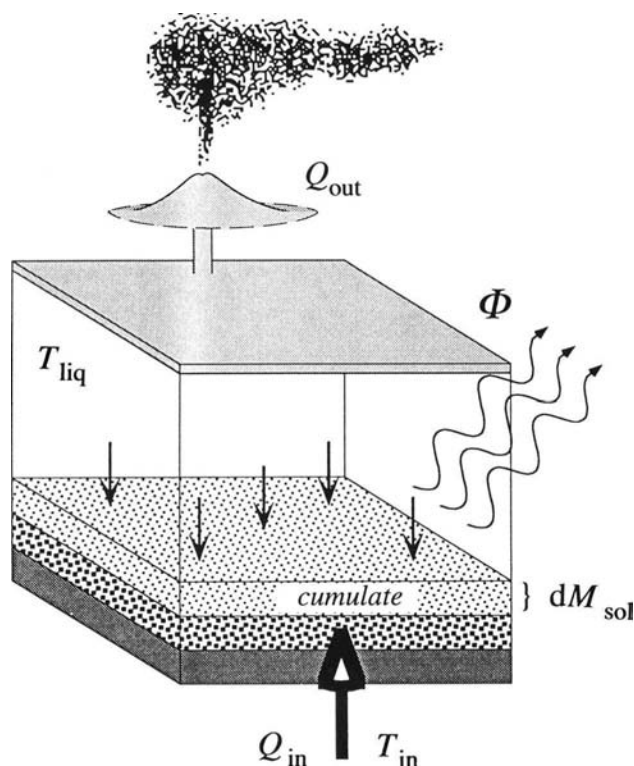


FIG. 1 A simple model of a magma chamber continuously erupting, crystallizing, and replenished. Cooling of the liquid reservoir by conductive or hydrothermal heat loss to the surrounding rocks triggers thermal runaway which in turn results in chemical discontinuities. M_{sol} refers to the mass of cumulate, Q to mass fluxes, and Φ to the heat flux to the surroundings. Continuous replenishment and cumulate sedimentation are accommodated by a displacement of the wall rocks.

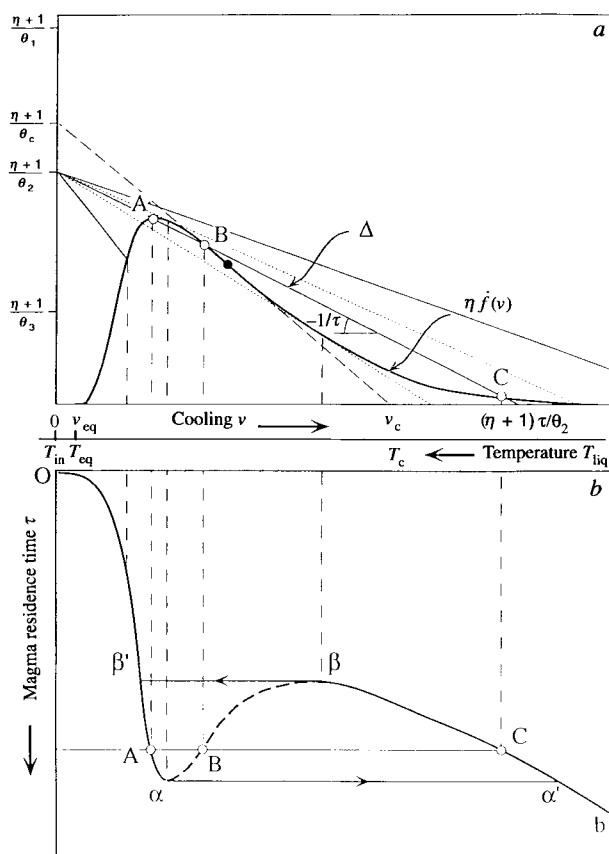


FIG. 2 Stability analysis of the magmatic system represented in Fig. 1. In a, v is the dimensionless cooling, given by $1 - T_{liq}/T_{in}$. Magma residence time is τ , newtonian cooling time is θ , rate of crystallization is f , and reduced latent heat is η . Thermal steady state occurs at any intersection of the bell-shaped curve $\eta f(v)$ and the straight line Δ . The thermal evolution is conditioned by the rate of cooling and the magma residence time. Critical conditions are those defined by the tangent to the curve $\eta f(v)$ at its inflection point (filled circle). For $\theta = \theta_1 > \theta_c$ (slow cooling), only one steady state is possible. For $\theta = \theta_2 < \theta_c$ two or three steady states become accessible to the system. The system bifurcates when Δ becomes tangent to $\eta f(v)$. b, Stable (solid line) and unstable (dashed line) branches of the steady-state locus for $\theta = \theta_2 < \theta_c$. Decreasing the rate of replenishment of the reservoir by fresh magma makes Δ rotate anti-clockwise in a. The system cools at steady-state down the branch $O\alpha$. Then magma abruptly chills along $\alpha\alpha'$ where progressive cooling may continue along the branch $\alpha'b$ (magmatic catastrophe). Note that the path is not reversible (hysteresis).

bifurcation, especially if cooling of the reservoir is enhanced by groundwater circulation.

The parametrization of eruptive dynamics is expectedly far more complicated for hydrous magmas. However, a thermal bifurcation induced by an apparently small change in the rate of replenishment will result in massive crystallization, which might leave volatile-saturated residual magmas and in rapid outgassing that may affect the eruptive regime of the magmatic reservoir. Although the case for a heterogeneous magma chamber violating the premises of the present model is much stronger for wet felsic systems, evidence of magma residence times as long as 10^5 – 10^6 years in such a very active system as the Long Valley caldera^{20,21} hints at the possibility of thermal catastrophes associated with ash-flow eruptions.

Some limitations apply to the present model. It has initially been assumed that the lava removed is aphyric. The formalism developed above can be extended to phryic lavas provided the relative rate of cumulate and magma (=liquid + suspended crystals) withdrawal maintains the relative proportion of suspended

solid and liquid in the magma chamber, which is a condition for steady-state behaviour. A reservoir holding a variable volume of magma could be accounted for by a more detailed analysis. Finally, coupled thermal and chemical dynamics makes the problem substantially more complicated^{8,22}.

Vigorous hydrothermal cooling of intrusions, such as that induced by pervasive seawater circulation in mid-ocean ridges, should also favour magmatic bifurcation. Petrological work on ODP leg 118²³ has revealed a quite spectacular and consistent magmatic gap between olivine gabbros and more differentiated oxide gabbros which are thought to be genetically related. Oxide gabbros are interpreted as late-stage differentiates injected into olivine gabbro at various levels along shear zones. A puzzling aspect of this association is the striking consistency of olivine, clinopyroxene and plagioclase compositions in each rock type, and the lack of intermediate mineral compositions in sequences that repeat themselves many times over a few hundred metres and which we suggest reflect a hydrothermally induced magmatic bifurcation. □

Received 26 May; accepted 3 October 1995.

- Daly, R. A. *Proc. Am. Acad. Arts. Sci.* **60**, 1–80 (1925).
- Chayes, F. J. *geophys. Res.* **68**, 1519–1534 (1963).
- Clague, D. A. *J. Geol.* **86**, 739–743 (1978).
- Jones, W. B. *Geol. Mag.* **116**, 487–489 (1979).
- Baker, B. H., Goles, G. G., Leeman, W. P. & Lindstrom, M. M. *Contr. Miner. Petrol.* **60**, 303–322 (1977).
- Brophy, J. G. *Contr. Miner. Petrol.* **109**, 173–182 (1991).
- Marsh, B. D. *Contr. Miner. Petrol.* **78**, 85–98 (1981).
- Gray, P. & Scott, S. K. *Chemical Oscillations and Instabilities* (Oxford Univ. Press, 1994).
- Dowty, E. in *Physics of the Magmatic Processes* (ed. Hargraves, H. B.) 419–485 (Princeton Univ. Press, 1980).
- Tiller, W. A. *The Science of Crystallization: Macroscopic Phenomenon and Defect Generation* (Cambridge Univ. Press, 1991).
- Tiller, W. A. *The Science of Crystallization: Microscopic Interfacial Phenomena* (Cambridge Univ. Press, 1991).

- Klein, L. C. & Uhlmann, D. R. *Proc. Lunar Sci. Conf.* **7**, 1113–1121 (1976).
- Brandeis, G. & Jaupart, C. *Contr. Miner. Petrol.* **96**, 24–34 (1987).
- Mangan, M. T. *J. Volcan. geotherm. Res.* **44**, 295–302 (1990).
- Albarède, F. *Geochim. cosmochim. Acta* **57**, 615–621 (1993).
- Condomines, M., Tanguy, J.-C. & Michaud, V. *Earth planet. Sci. Lett.* **132**, 25–41 (1995).
- Clague, D. A. *Bull. volcan.* **49**, 577–587 (1987).
- Cortini, M. & Scandone, R. *J. volcan. geotherm. Res.* **12**, 393–400 (1982).
- Carlsaw, H. S. & Jaeger, J. C. *Conduction of Heat in Solids* (Oxford Univ. Press, 1959).
- Halliday, A. N. et al. *Earth planet. Sci. Lett.* **94**, 274–290 (1989).
- Christensen, J. N. & DePaolo, D. J. *Contr. Miner. Petrol.* **113**, 100–114 (1993).
- Bonnefoi, C. *Modélisation de la Dynamique Géochimique des Chambres Magmatiques* (DEA Memoir, Clermont-Ferrand, 1994).
- Dick, H. J. B., Robinson, P. T. & Meyer, P. S. in *Synthesis of Results from Scientific Drilling in the Indian Ocean 1–39* (Geophys. Monogr. 70, Am. Geophys. Union, Washington DC, 1992).

ACKNOWLEDGEMENTS. We thank J. Blichert-Toft and W. M. White for comments on the manuscript. This work is a contribution from URA 726 and URA 10 (CNRS).

The first australopithecine 2,500 kilometres west of the Rift Valley (Chad)

Michel Brunet*, Alain Beauvilain†, Yves Coppens‡, Emile Heintz§, Aladji H. E. Moutaye|| & David Pilbeam¶

* Laboratoire Géobiologie, Biochronologie et Paléontologie humaine, JE 273 MST Faculté des Sciences de l'Université de Poitiers, 86022 Poitiers Cedex, France

† Centre National d'Appui à la Recherche (CNAR), BP 1228, N'Djamena, Chad

‡ Chaire de Paléanthropologie et Préhistoire, Collège de France, 75005 Paris, France

§ CNRS, 3 rue Michel Ange, 75794 Paris Cedex 16, France

|| Direction de Recherches Géologiques et Minières (DRGM), N'Djamena, Chad

¶ Department of Anthropology, Harvard University, Cambridge, Massachusetts 02138, USA

THE first sites with Pliocene and Pleistocene mammals west of the Rift Valley in Central Africa in northern Chad were reported in 1959 (ref. 1), and documented the presence of mixed savannah and woodland habitats. Further sites² and a probable *Homo erectus* cranio-facial fragment³ were subsequently discovered. In 1993 a survey of Pliocene and Pleistocene formations in the Borkou-Ennedi-Tibesti Province of Chad (B.E.T.) led to the discovery of 17 new sites in the region of Bahr el Ghazal (classical Arabic for River of the Gazelles) near Koro Toro. One site, KT 12 (15°58'10" N, 18°52'46" E) yielded an australopithecine mandible associated with a fauna biochronologically estimated to be 3.0–3.5 Myr old. Australopithecine species described since 1925 are known from southern Africa and from sites spread along the eastern Rift Valley from Tanzania to Ethiopia (Fig. 1). This new find from Chad, which is most similar in morphology to *Australopithecus afarensis*⁴, documents the presence of an early hominid a considerable distance, 2,500 km, west of the Rift Valley.

In the Chad basin, Pliocene sediments are generally covered by Quaternary ones and are known mostly from boreholes. However, they are exposed around Koro Toro, albeit rather poorly. Two sets of sediments are normally recognized, a lower sandy level lying in disconformity on terrestrial deposits (Continental Terminal) and in part originating from these deposits, and an upper level of clays with sandy intercalations^{5,6}. Features of the KT 12 fluvio-lacustrine facies indicate reworking of older sediments (rounded quartz, perhaps from the Continental Terminal), and suggest a very low-energy fluvial system (angular quartz and no graded bedding). The latter suggestion is supported by the discovery of an articulated *Ceratotherium* skeleton.

The fossil vertebrates from KT 12 consist of surface finds and some *in situ* material, both from the same level of poorly consolidated fine-grained sandstone. This level contains abundant quartz (some rounded, some angular) and some potash feldspar (microcline). Traces of hydrated sodium silicates indicate phases of evaporation during dry episodes. There is no graded bedding. Quartz with undulatory extinction indicates tectonic deformation. Other strongly corroded quartz indicates the presence of dissolution channels.

Preliminary identifications of the taxa from KT 12 are as follows.

Pisces: large Siluridae indet.

Reptilia

Chelonina: *Geochelone* sp.; *Trionyx* sp.

Crocodylia: cf. *Tomistoma* sp.

Mammalia

Proboscidea: *Loxodonta exoptata* as from Laetoli and Hadar^{7,8}.

TABLE 1 Dental measurements of *Australopithecus* aff. *A. afarensis* (KT 12/H1)

	I ₂	C		P ₃		P ₄	
	R	R	L	R	L	R	L
Mesiodistal (length) (mm)	5.5	9.0	9.0	9.7	9.0	10.3	9.5
Labiolingual (breadth) (mm)	7.6	11.0	11.0	12.5	12.7	12.1	12.5

Measurements are of the crowns of the preserved lower dentition. R, right; L, left.

Perissodactyla: *Ceratotherium* cf. *praecox*, as from Hadar⁹; *Hipparion* sp. aff. *H. afarensis*/H. *hasumense* as from Hadar^{10,11}.

Artiodactyla:

Giraffidae: *Sivatherium* sp.

Bovidae

Reduncini: *Kobus* sp.

Alcelaphini indet.

Bovini indet.

Suidae: *Kolpochoerus afarensis* as from Hadar and Laetoli^{12,14}.

Hippopotamidae: cf. *Hexaprotodon protamphibius turkanensis*, as from Omo^{15,16}.

Preliminary analysis suggests that the fauna from KT 12 shows closest resemblances to collections from Hadar with an approximate age of 3.0–3.4 Myr. The non-hominid fauna contains aquatic taxa (such as Siluridae, *Trionyx*, cf. *Tomistoma*), taxa adapted to wooded habitats (such as *Loxodonta*, *Kobus*, *Kolpochoerus*) and to more open areas (such as *Ceratotherium*, *Hipparion*). Assuming that there has not been extensive transport, as suggested by the sedimentary environment, we tentatively propose that the palaeontological and sedimentological evidence is compatible with a lakeside environment, with both perennial and permanent streams, and a vegetational mosaic of gallery forest and wooded savannah with open grassy patches.



FIG. 1 Geographic location of the major sites with primitive hominids: 1, Taung; 2, Kromdraai, Sterkfontein, Swartkrans; 3, Makapansgat; 4, Laetoli; 5, Olduvai; 6, Kanapoi, Lothagam; 7, Koobi Fora; 8, Omo; 9, Middle Awash; 10, Hadar; 11, Bahr el Ghazal (KT 12).

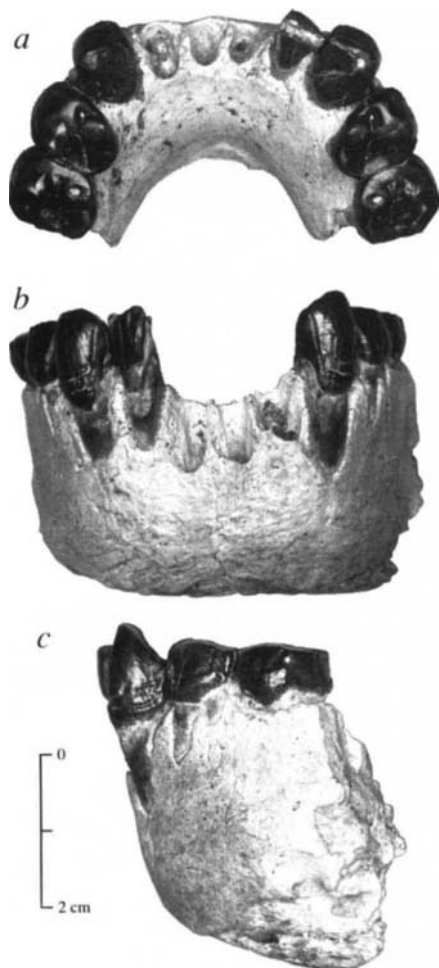


FIG. 2 *Australopithecus* aff. *A. afarensis* mandible (KT 12/H1) with the crowns of left P_3 – C_1 , alveoli of left I_2 , I_1 and right I_1 and the crowns of right I_2 – P_4 . a, Occlusal view; b, anterior view; c, left lateral view.

The new hominid discovery, KT 12/H1, is a perisymphseal fragment of an adult mandible with symmetrical fractures passing distal to the P_4 roots. It includes the crowns of the right I_2 – P_4 and the left C – P_4 (Figs 2 and 3, Table 1). The specimen is currently being studied at the University of Poitiers Faculty of Sciences, and will be permanently curated at CNAR in N'Djamena, Chad.

The specimen is most similar to *A. afarensis*⁴, with large incisiform canines, strongly bicuspid P_3 s, P_4 s with a modest talonid, and cingula on canines and premolars. It differs from *A. afarensis* (and other australopithecine specimens known to us) in several features, including the form of the lingual surface of the symphysis which has a subvertical planum alveolare, with small superior and inferior transverse tori, both premolars with three completely distinct roots, and premolar occlusal enamel that appears in computed axial tomography (CAT) scans to be thin. We will make extensive comparisons of the new specimen with the relevant hominid material from sites elsewhere in Africa. For the time being we refer the new specimen to *Australopithecus* aff. *A. afarensis*.

The discovery of an australopithecine mandible together with a middle Pliocene fauna 2,500 km west of the Rift Valley considerably extends the known range of these early hominids and raises several interesting issues. The Chad specimen is most similar to its East African contemporary *A. afarensis*. Nevertheless, in certain features—mandibular morphology, premolar roots and enamel thickness—it differs from the described hypodigm

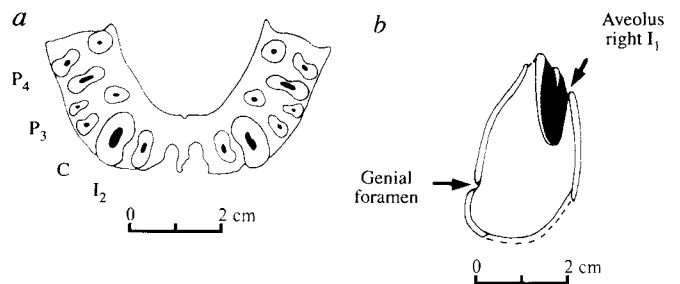


FIG. 3 Diagrams of KT 12/H1, based on CAT scans made with a Picker PQ 2000Z Scanner, with necessary adjustments made for differences in radio-opacity between dentine and enamel. a, Transverse section through the alveolar processes of mandible; b, section through mandibular symphysis.

of *A. afarensis*⁴. Given the genetic and morphological differences now recognized between allopatric populations within, for example, *Pan troglodytes*, *Gorilla gorilla*, *Pongo pygmaeus* and *Papio hamadryas*^{17–22} as well as other African mammals²³, it is not surprising that contemporaneous hominid populations as geographically distant as Chad and Ethiopia, Kenya and Tanzania would differ in morphology, regardless of whether they are classified as species or subspecies. Here we do not choose to name a new species, recognizing that more detailed comparisons are necessary before the taxonomy of this Bahr el Ghazal hominid can be resolved.

Eastern and southern Africa, the regions from which the seven or so species of australopithecine have been currently described^{24,25}, are separated by more than 2,000 km; similar but different species are present in the two regions, and related hominids surely inhabited the intermediate areas. No apes have been collected throughout this entire zone. Over the past 20 years^{26–28} there has been a growing consensus that the western Rift Valley, which developed during the late Miocene, played a major role in the origin of hominids, segregating more forested chimpanzee habitats to the west from more open hominid habitats to the east, with eastern Africa seen as being particularly significant.

The presence of middle Pliocene hominids in Chad, a further 2,500 km to the west of the western Rift Valley, suggests that at least by that time hominids were distributed throughout the woodland and savannah belt from the Atlantic Ocean across the Sahel through eastern Africa to the Cape of Good Hope. The early Pliocene and probably late Miocene presence of hominids in eastern Africa²⁴ supports an origin there for the hominids, but this inference is based entirely on the fact that comparably aged sediments are either absent from or have not yet been located in southern, central and western Africa. If the origins of hominids occurred rapidly, followed by rapid range extension, as seems likely, it may be as futile to seek a specific and localized place of origin for hominids as it is for any other group. □

Received 30 June; accepted 26 September 1995

1. Abadie, J., Barbeau, J. & Coppens, Y. *C. r. Acad. Sci., Paris* **248**, 3329–3330 (1959).
2. Coppens, Y. in *Background to Evolution in Africa* (eds Bishop, W. W. & Clark, J. D.) 89–97 (Univ. Chicago Press, 1967).
3. Coppens, Y. *C. r. Acad. Sci., Paris* **260**, 2869–2871 (1965).
4. Johanson, D. C., White, T. D. & Coppens, Y. *Am. J. phys. Anthropol.* **57**, 545–603 (1982).
5. Schneider, J. L. thesis, Univ. Avignon (1989).
6. Schneider, J. L. & Wolff, J. P. *Document du BRGM, Orléans* **209**, (1992).
7. Beden, M. in *L'environnement des Hominidés au Plio-Pleistocène* (ed. Fondation Singer Polignac) 21–44 (Masson, Paris, 1985).
8. Leakey, M. D. et al. *Nature* **262**, 460–466 (1976).
9. Guérin, C. *Bull. Soc. géol. Fr.* **XXI**, 283–288 (1979).
10. Johanson, D. C. in *Koobi Fora Project* (ed. Harris, J. M.) 156–214 (Clarendon, Oxford, 1993).
11. Eisenmann, V. *Geobios* **9**, 577–605 (1976).
12. Harris, J. M. & White, T. D. *Trans. Am. phil. Soc.* **69**, 1–128 (1979).
13. Harris, J. M. in *Koobi Fora Project* (ed. Harris, J. M.) 215–302 (Clarendon, Oxford, 1983).
14. Cooke, H. B. S. in *L'environnement des Hominidés au Plio-Pleistocène* (ed. Fondation Singer Polignac) 101–117 (Masson, Paris, 1985).
15. Geze, R. thesis, Univ. Paris (1980).

16. Geze, R. in *L'environnement des Hominidés au Plio-Pleistocène* (ed. Fondation Singer Polignac) 81–100 (Masson, Paris, 1985).
17. Ruvoilo, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 8900–8904 (1994).
18. Morin, P. A. et al. *Science* **265**, 1193–1201 (1994).
19. Uchida, A. *Am. J. phys. Anthropol.* (suppl.) **16**, 198–199 (1993).
20. Groves, C. P. in *Comparative Primate Biology: Systematics, Evolution and Anatomy* (eds Swindler, D. R. & Erwin, J.) 187–217 (Liss, New York, 1986).
21. Braga, J. thesis, Univ. Bordeaux (1995).
22. Jolly, C. J. in *Species, Species Concepts, and Primate Evolution* (eds Kimbel, W. H. & Martin, L. B.) 67–108 (Plenum, New York, 1993).
23. Grubb, P. in *Pigs, Peccaries and Hippos Status Survey and Conservation Action Plan* (ed. Oliver, W. L. R.) 66–75 (IUCN, Gland, Switzerland, 1993).
24. Wood, B. A. in *The Cambridge Encyclopedia of Human Evolution* (eds Jones, S., Martin, R. & Pilbeam, D.) 231–240 (Cambridge Univ. Press, 1992).
25. White, T. D., Suwa, G. & Asfaw, B. *Nature* **371**, 306–312 (1994).
26. Kortlandt, A. *New Perspectives on Ape and Human Evolution*, 9–100 (Stichting voor Psychobiologie, Amsterdam, 1972).
27. Coppens, Y. *Le singe, l'Afrique et l'Homme* (ed. Jacob, O.) 1–148 (Fayard, Paris, 1983).
28. Pickford, M. *Hum. Evol.* **5**, 1–20 (1990).

ACKNOWLEDGEMENTS. We informally name this primitive hominid 'Abel', in dedication to the memory of our colleague and friend Abel Brillanceau who died in the field in Cameroon on 27 February 1989. This research was made possible by the Ministère Tchadien de l'Éducation Nationale and the Centre National d'Appui à la Recherche (CNAR); Ministère Tchadien des Mines et de l'Énergie and the Direction de Recherches Géologiques et Minières (DRGM); Programme des Nations Unies pour le développement (PNUD); Mission française d'Aide et de Coopération; and the Faculté des Sciences de l'Université de Poitiers. We thank the French Army Général Kréher Conseiller Militaire Spécial, N'Djamena and the MAM of Moussoro (4e RE et 4e RIMA) and Faya; the other expedition members, N. Beauvilain and M. Tormalita; P. Chartier for providing access to Scanner-IRM Poitou-Charentes; P. Fronty and M. Sapanet for radiology interpretations; J. P. Bassouillet, E. Cariou, A. Decarreau and J. J. Trescases for sedimentologic environment; D. Geraads and L. de Bonis for assistance; V. Bellefleur for photography; G. Florent for secretarial support; G. Mouchelin for the cast; and S. Riffaut for the drawings.

Early *Homo* and associated artefacts from Asia

Huang Wanpo*, Russell Ciochon†‡, Gu Yumin*, Roy Larick§, Fang Qiren||, Henry Schwarcz¶, Charles Yonge*, John de Vos* & William Rink¶

* Institute of Vertebrate Paleontology and Paleoanthropology, Academia Sinica, Beijing 100044, China

† Departments of Anthropology and Pediatric Dentistry, The University of Iowa, Iowa City, Iowa 52242, USA

§ Department of Anthropology, University of Massachusetts, Amherst, Massachusetts 01003, USA

|| Chongqing Museum of Natural History, 72 Loquat Hill Park Street, Chongqing, Sichuan 630013, China

¶ Department of Geology, McMaster University, Hamilton, Ontario L8S 4M1, Canada

* Department of Physics & Astronomy, University of Calgary, Calgary, Alberta T2N 1N4, Canada

¶ National Museum of Natural History, Post Office Box 9517, 2300 RA Leiden, The Netherlands

THE site of Longgupo Cave was discovered in 1984 and excavated in 1985–1988 by the Institute of Vertebrate Paleontology and Paleoanthropology (Beijing) and the Chongqing National Museum (Sichuan Province). Important finds include very archaic hominid dental fragments, *Gigantopithecus* teeth and primitive stone tools. Palaeomagnetic analysis and the presence of *Ailuropoda microta* (pygmy giant panda) suggested that the hominid-bearing levels dated to the earliest Pleistocene¹. In 1992, joint Chinese–American–Canadian geochronological research corroborated the age using electron spin resonance (ESR) analysis. We report here that the hominid dentition and stone tools from Longgupo Cave are comparable in age and morphology with early representatives of the genus *Homo* (*H. habilis* and *H. ergaster*) and the Oldowan technology in East Africa. The Longgupo dentition is demonstrably more primitive than that seen in Asian *Homo erectus*. Longgupo's diverse and well preserved Plio-Pleistocene fauna of 116 species provide a sensitive contextual base for interpreting the early arrival of the genus *Homo* in Asia.

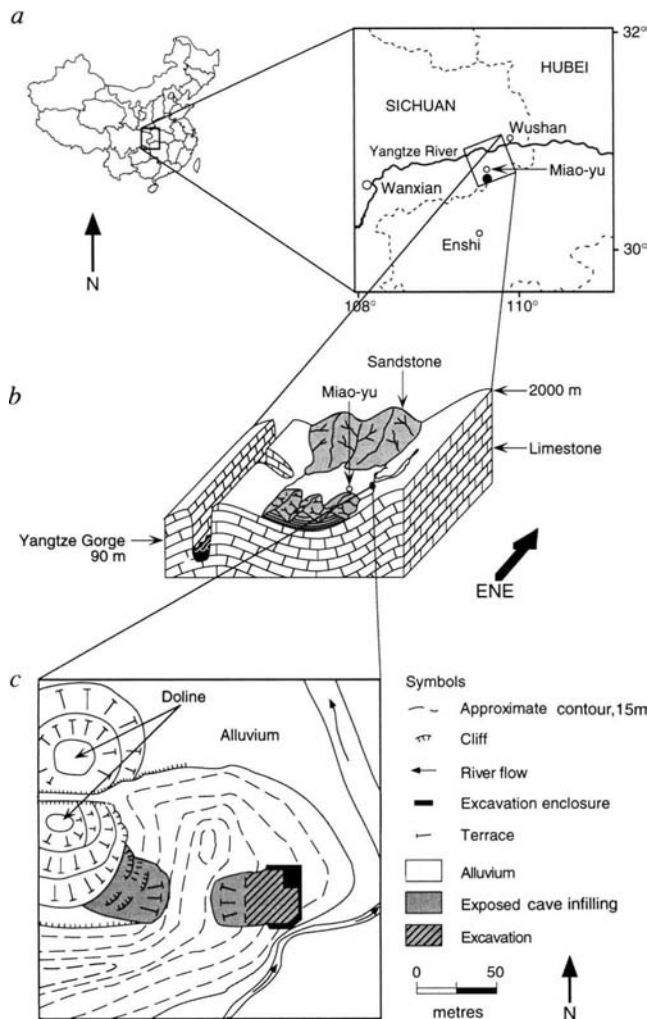
‡ To whom correspondence should be addressed.

Longgupo Cave or, as known in Chinese, the 'Wushan Hominid Site' lies 20 km south of the Yangtze River near the eastern border of Sichuan Province (Fig. 1). The cave infilling comprises two major units: an upper cave breccia with few fossils, and an underlying poorly cemented fossiliferous unit with three depositional zones (Fig. 2). The middle zone (excavation levels 2–12) has a clay facies localized along the north and south cave walls in longitudinal channels 2–3 m wide. Two hominid dental fragments and two stone artefacts derive from these channels. The middle zone yields 68 mammalian genera including *Procynocephalus* and *Macaca* as well as *Gigantopithecus* and *Homo*¹. The presence of *Sinomastodon*, *Nestoritherium*, *Equus yunnanensis*, *Ailuropoda microta* and Cricetinae indeterminate (with molar structure similar to *Sinocricetus* and *Nannocricetus*) suggests a late Pliocene to earliest Pleistocene age^{1–3}. The occurrence of *Mimomys peii* places this zone within the Dachaian (mammal neogene reference level 17 (MN 17)) age of north China (late Pliocene)^{4,5}.

Basic taphonomic observations suggest how the fauna accumulated within the middle zone channels. About 750 large mammal long bones derive from the channels, usually without surface erosion and sometimes as complete limb elements. Many shafts are punctured, split or shattered to indicate processing by *Pachycrocuta* or *Homotherium*, and many long bones were also gnawed, probably by *Hystrix* and smaller rodents while fresh. *Pachycrocuta* is the most common carnivore at Longgupo, and numerous hyaenid-sized coprolites suggest the cave occasionally served as a lair. Two accumulation processes are therefore possible. The large scavengers/carnivores may have brought carcasses into a horizontal passage, eventually to be interred by gentle fluvial action. Alternatively, the representation and condition of the Longgupo fossils also recall the assemblages from Yanjinggou (about 50 km NW), where carcasses accumulated in vertical passages, the result of predation and falls^{6,7}.

Palaeomagnetic data also suggest great antiquity for the middle zone clay facies where 100 sediment samples were taken in overlapping columns. Sample natural remanent magnetization was determined after thermal demagnetization to 700 °C. Sediment at the top of the lower unit is magnetically reversed, indicating an age of at least 0.78 million years (Myr). There are 7 magnetic reversals between levels 1 and 20 with magnetic normal sediment sections thinner (about 1–2 m) than reversed sections (about 3–4 m) (Fig. 2). Thermal demagnetization properties are similar in normal and reversed sediments, suggesting both remanence types are of the same (detrital) origin. Although we do not have sufficient rock magnetic data to exclude completely the possibility that selective remagnetization may have led to a spurious magnetostratigraphy, we do not believe that this is likely. We therefore assign the magnetically normal hominid-bearing levels 7–8 to the Olduvai event that spans the interval 1.96 to 1.78 Myr^{8,9}. A few depositional breaks (such as at the 13/12 boundary) may correspond to lengthy time gaps in which further reversals could have occurred. Therefore this is a minimum age estimate for the hominids.

Electron spin resonance (ESR) analysis of tooth enamel further constrains the age of the middle zone channels. This method has been successfully applied to fossils older than 2 Myr^{10,11}. From level 4, a large cervid premolar fragment (recovered within the sediment-filled interior of a long bone) yields about 400 mg of enamel and dentine, containing 0.85 and 39 p.p.m. uranium, respectively. External gamma dose-rates were estimated from U, Th and K contents of sediment adhering outside the long bone; beta dose-rates alone were determined from sediment within the long bone cavity. Age calculation depends on an assumed uranium uptake history of the tooth: early uptake gives a minimum age of 0.75 ± 0.09 Myr; a linear uptake model, which generally has given ages close to independent estimates, yields 1.02 ± 0.12 Myr¹¹. The linear uptake age places the level 4 sediment within the Matuyama reversed mag-

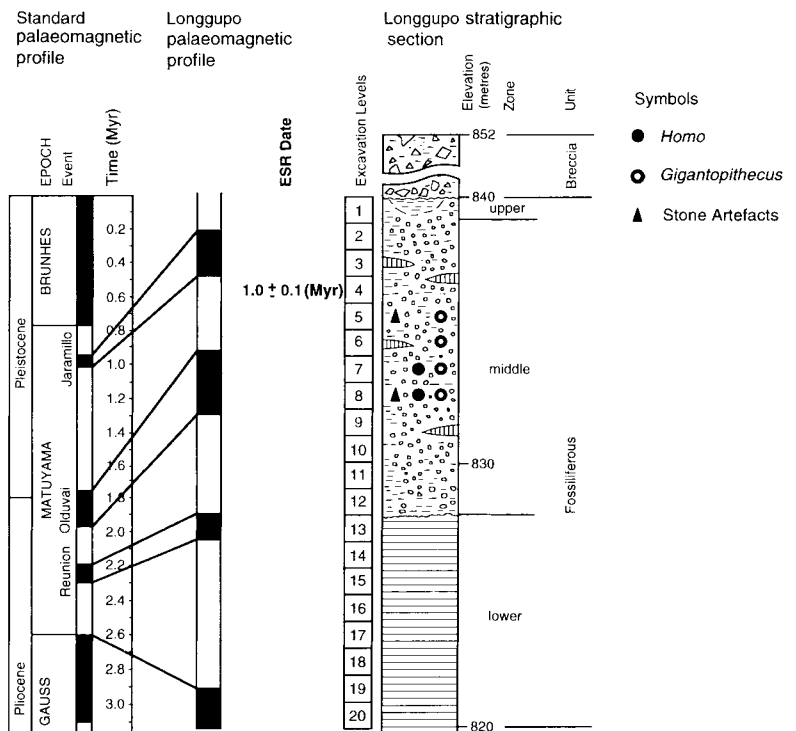


netic chron, consistent with the palaeomagnetic interpretation for levels 7–8.

Within the middle zone channels, levels 7–8 yield two hominid dental fossils, a fragmentary left mandible (with P_4 , M_1 and M_2 alveolus) and a right upper lateral incisor (Fig. 3). These specimens were originally described as a new subspecies of *Homo erectus*, *H. e. wushanensis*¹, but joint reanalysis demonstrates morphology different from and sometimes more primitive than Asian *H. erectus*. The Longgupo dentition exhibits affinities with representatives of East African early Pleistocene *Homo*, including *H. ergaster* and *H. habilis* (Fig. 3 legend). For the mandibular teeth, relevant features include, for P_4 : cusp placement and inflection, relationship between cusps and talonid, and root morphology; and for M_1 : cusp number and spatial relationship, enamel surface thickness, and surface texture. Basic dimensions of the Longgupo mandibular corpus, P_4 and M_1 (Fig. 3 legend) are very small. These fall outside the lower range for Zhoukoudian (G1, H1, 29, 30, 34, 89, 93, 96–102)¹² and Sangiran (1b, 9) specimens^{13,14}, while lying near the lower limit for any East African Plio-Pleistocene hominine (such as OH 7, OH 13, OH 16, ER 992 and WT 15000)^{14,16}. I^2 crown shape index ($BL/LL \times 100$) also distinguishes Longgupo from Asian *H. erectus*. The Longgupo index (116) is quite high compared with Zhoukoudian 6 and 7 (101), but falls within the range for OH 6, OH 16, OH 39 and ER 1813 (*H. habilis*) (96–127 (mean = 109))¹⁵. Finally, the shovel-shaped character of the Longgupo

FIG. 1 Longgupo Cave. a, Location in Sichuan Province, Wushan County (109° 40' E, 30° 50' N). b, Regional geology has massive Triassic limestones and intercalated sandstones. The primary local geomorphic feature is an E–W-trending syncline that exposes carbonates on its limbs and silicates within. The limbs are karstic and the consequent drainage forms a polje around Miao-yu village. Longgupo Cave is a remnant of this drainage that has been eroding the polje's southern margin since Miocene times. c, The cave presents an infilled east–west passage, bi-truncated to leave its floor 130 m long and its vault 25 m long. Vertical passages are also common within the local complex. Just northwest of the Longgupo site, a lower north–south passage retains an even shorter vault between two dolines.

FIG. 2 Longgupo Cave stratigraphy and suggested age. The upper sedimentary unit is a 12-m-thick fossil-poor cave breccia where clasts range to 50 cm × 90 cm within a highly cemented sandy clay. The underlying poorly cemented fossiliferous unit has three depositional zones excavated in 1 m levels. The upper zone (within level 1) consists of a few spatially discrete sandy clay lenses with some gravels and localized areas of calcite concretion. The middle zone (levels 2–12) has a clay facies localized along the north and south cave walls in longitudinal channels (2–3 m wide), while a gravel facies occupies the passage centre. Clast size ranges to 2 cm × 3 cm in the clay facies and to 10 cm × 20 cm in the gravels with limestone clasts dominating shale and mudstone. The lower zone (levels 13–20) has primarily silts in parallel horizontal beds that indicate stagnant fluvial or lacustrine environments.



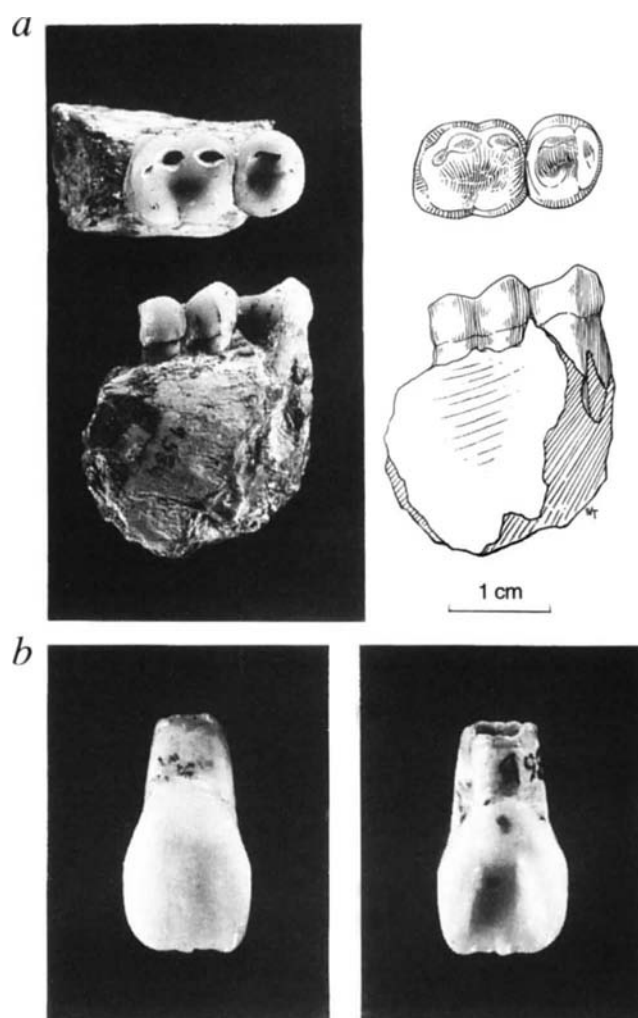
I² is also found in OH 6 and OH 16 (*H. habilis*)¹⁵, and WT 15000 (*H. ergaster*)¹⁶ as well as at Zhoukoudian¹². Indeed, Tobias concludes that "shovelling is a virtually universal feature of all early hominids up to *Homo erectus*" (p. 614)¹⁵.

Sixteen teeth of *Gigantopithecus blacki* come from levels 5–8 (Fig. 2): one C₁, three P₃, one P³, four P₄, three P⁴ and four M_{1,2}. In size and morphology the Longgupo teeth match the >900 isolated teeth from Liucheng Cave, Guangxi, China^{17,18}, where amino-acid racemization measurements suggest an age greater than 1.2 Myr¹⁹. The Longgupo *Gigantopithecus* teeth, like those of Liucheng, exhibit high sexual dimorphism, lower molars with a distinctive accessory internal tubercle between metaconid and entoconid and very thick enamel (5–6 mm). Nearly all of the roots show evidence of rodent gnawing, probably by *Hystrix*. Longgupo marks the northernmost extent of *Gigantopithecus* in China. It also represents the third east Asian cave where *Gigantopithecus* and *Homo* lie within the same stratigraphic interval. The others are Tham Khuyen Cave, Lang Son, Vietnam dated to 0.475 Myr²⁰, and at Jianshi Cave, Hubei,

China, where the co-occurrence remains undated²¹. With the Longgupo date indicating the earliest presently known association, *Gigantopithecus* and *Homo* evidently coexisted for more than 1 Myr in east Asia.

Two items from the clay facies show three characteristics of early stone artefacts (Fig. 4). They have exotic petrological composition (andesite–porphyrite), they are twice the size of the largest normal deposit clasts, and they exhibit surface damage inconsistent with natural cause. The closest primary outcrops for andesite–porphyrite rocks lie 130 km ENE and 150 km NNW. Derived sources occur within the non-carbonate sediments of the Miao-yu polje, a function of pre-Yangtze Valley (Pliocene and earlier) fluvial systems. Potential sources are to be found about 2 km downstream from Longgupo, but andesite–porphyrite remains undocumented. As both objects are more than twice as large as the largest clay facies clasts, fluvial transport does not explain their presence. Assuming artefact status, these specimens recall the Oldowan technology of early Pleistocene East Africa²² in two ways. First, they were chosen

FIG. 3 Longgupo Cave hominids. a, Left mandible fragment (CV.939.1, level 8), occlusal (top) and lingual views. The mandibular corpus is nearly complete directly below M₁, where the inferior margin is damaged. The buccal and lingual faces bulge little and remain parallel as they descend to the inferior margin. Behind M₁, the buccal face begins to thicken. Below M₁, corpus height is 21 mm and width is 13.5 mm. For the limited morphology preserved, the Longgupo mandibular corpus is gracile compared with Asian *H. erectus* (Zhoukoudian G1, H1; Sangiran 1b, 9)^{12–14}, and more closely resembles specimens of early Pleistocene East African *Homo* such as KNM-WT 15000, KNM-ER 730 (*H. ergaster*) and OH 13 (*H. habilis*)^{14–16}. P₄ (mesio-distal (MD)=7.4, buccal-lingual (BL)=9.1): The crown is slightly subcircular with a buccal-lingual long axis. No cingulum is present. The cusps are placed mesially and inflected centrally to be separated only by a central sagittal groove, while a long, wide talonid, with no cusps or ridges, occupies the distal two-thirds of the tooth. The neck is constricted, and a long robust root bifurcates for 2/3 of its length. In comparison, the seven Asian *H. erectus* P₄ from Zhoukoudian (29, 30, 89–93) have an obliquely oval crown with the greatest diameter running mesio-buccally to distal-lingually, with distinct buccal and lingual cusps, and with a small talonid having numerous short accessory ridges and wrinkles. The buccal surface of these teeth has a well developed cingulum and the root structure shows a single tapering root that is compressed mesio-distally. In the Longgupo P₄, the buccal-lingually expanded crown resembles ER 992 (*H. ergaster*), the position of the mesial cusps and large talonid compares with OH 7 and OH 13 (*H. habilis*), and the bifid root is like ER 992 (*H. ergaster*) and to a lesser extent OH 13 (*H. habilis*). M₁ (MD=11.0, BL=10.1): The tooth is low-crowned and possesses two strong roots and five cusps. In occlusal view the crown is sub-rectangular with rounded corners. The metaconid is larger than entoconid and is the highest cusp. The enamel is relatively thin and uncrenulated; perforations expose dentine at the protoconid, hypoconid and hypoconulid. In comparison, the seven *in situ* M₁ of *Homo erectus*, as well as Black's type specimen (Zhoukoudian 96–102, 34)¹² exhibit six cusps (including tubercle 6), thickened enamel, cusps covered by wrinkles, furrows and accessory ridges, and larger size. The cusp arrangement and size of the Longgupo M₁ may be likened with ER 992 (*H. ergaster*) and OH 13 (*H. habilis*). b, Right I² (CV.939.2, level 7) labial (left) and lingual views (×1.5 the scale of a). This isolated tooth is complete and unerupted with the root intact (MD=8.1, labio-lingual (LL)=7.0, crown height (CH)=10.3). In lingual view, the mesial and distal marginal ridges are both high and well defined. A small tubercle is present at the base of the lingual surface where the mesial and distal marginal ridges converge upward to meet. The lingual surface is concave and quite shovel-shaped. A central ridge runs from the tubercle to the occlusal edge and divides into two branches at the middle of its crown. The labial surface is moderately convex. A mamelonated incisal ridge shows that the tooth was never in occlusion. The Longgupo I² differs from Zhoukoudian 6 and 7 (ref. 12) in its significantly lower crown height, less buccal-lingual expansion of the crown, significantly thicker mesial and distal marginal ridges, and a less wedge-shaped appearance when viewed from the lingual side of the occlusal surface. These features are variably expressed in the OH 6, 16, 39, ER 1813 (*H. habilis*)¹⁵ and WT 15000 (*H. ergaster*)¹⁶.



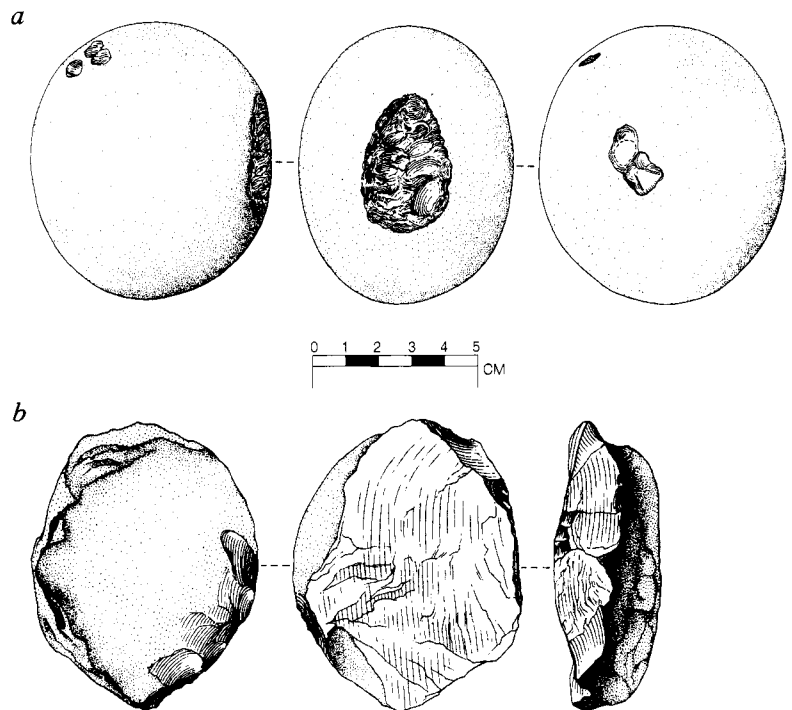


FIG. 4 Longgupo Cave andesite/porphyrite stone artefacts. *a*, The elongated, spherical cobble (P.6524, level 5) has a heavily weathered cortical surface. Three relatively discrete areas (left, centre and right) have reweathered crushing and pitting. The area at centre has a concentrated pattern of pits (4 mm depth) that suggest repeated battering. *b*, The lenticular flake (P.6523, level 8) represents a larger cobble that has split near a natural waist. The heavily abraded and weathered ventral surface consists of a few large scars (centre). The dorsal surface (left) shows two stages of weathering: rough, highly pitted areas confined to recesses, and smoother crests and convexities covering most of the surface. The dorsal surface recurves onto the ventral surface to indicate the waist in the original cobble (right). Heavily weathered crest damage occurs (left) clockwise from bottom centre at 0° – 85° and at 180° – 270° . The dorsal faceting at 270° – 350° also has heavy crest damage.

from the natural environment for inherent qualities of raw material and morphology and thus used with little modification²³. The andesite–porphyrite raw material is resilient to repeated battering, and both implements fit within the hand to offer numerous surfaces and edges for use. Second, the tightly patterned distribution of surface and edge damage indicates deliberate and consistent gestures of use²⁴. If, eventually, andesite–porphyrite cobbles prove unavailable within several kilometres of the Longgupo cave site, curation may distinguish this potentially earliest Asian technology from its African contemporaries.

We have estimated the age of the Longgupo hominid specimens and artefacts in three ways. The co-occurrence of *Sinomas-todon*, *Nestoritherium*, *Equus yunnanensis*, *Ailuropoda microta* and *Miomys peii* place the middle zone clay facies in the late Pliocene and earliest Pleistocene. Palaeomagnetic data then bracket levels 7–8 within the Olduvai chron, 1.96 to 1.78 Myr. Finally, the ESR determination of 1.02 ± 0.1 Myr, on material from level 4 (3 m above levels 7–8), constrains and reinforces these interpretations. The Longgupo specimens are therefore

older than any reported for China and at least as old as Indonesian *Homo erectus*²⁵. The Longgupo specimens closely resemble East African fossils representing the earliest species of the genus *Homo*. They share few characters in common with Asian *Homo erectus*. As their incompleteness precludes designating a new species, we assign the Longgupo hominids to *Homo* species indeterminate, while noting affinities with *H. habilis* and *H. ergaster*. The stone tools are consistent with this interpretation. Given the early date and primitive morphology for the Longgupo specimens, and the older age estimates for *H. erectus* in Java²⁵, we must recognize more than one Plio-Pleistocene hominid species in east Asia. The new evidence suggests that hominids entered Asia before 2 Myr, coincident with the earliest diversification of genus *Homo* in Africa^{26,27}. Clearly, the first hominid to arrive in Asia was a species other than true *H. erectus*, and one that possessed a stone-based technology. A pre-*erectus* hominid in China as early as 1.9 Myr provides the most likely antecedents for the *in situ* evolution of *Homo erectus* in Asia^{27–29}. □

Received 12 May; accepted 20 September 1995.

- Huang, W. et al. *Wushan Hominid Site* (Ocean, Beijing, 1991).
- Wang, L. & Ouyang, L. *Vertebr. Palasiatica* **20**, 255–263 (1982).
- Huang, W. *Vertebr. Palasiatica* **31**, 191–207 (1993).
- Zheng, S. & Li, C. *Vertebr. Palasiatica* **24**, 81–109 (1986).
- Zheng, S. *Quaternary Rodents of Sichuan-Guizhou Area, China* (Beijing, 1993).
- Matthew, W. D. & Granger, W. *Bull. Am. Mus. nat. Hist.* **48**, 563–598 (1923).
- Colbert, E. D. & Hooijer, D. A. *Bull. Am. Mus. nat. Hist.* **102**, 1–134 (1953).
- Cande, S. C. & Kent, D. V. *J. Geophys. Res.* **97**, 13917–13951 (1992).
- McDougall, I., Brown, F. H., Cerling, T. E. & Hillhouse, J. W. *Geophys. Res. Lett.* **19**, 2349–2352 (1992).
- Grün, R., Schwarcz, H. P. & Zymela, S. *Can. J. Earth Sci.* **24**, 1022–1037 (1987).
- Schwarcz, H. P., Grün, R. G. & Tobias, P. V. *J. hum. Evol.* **26**, 175–181 (1994).
- Weidenreich, F. *Paleontol. Sinica* **D1**, 1–180 (text); 1–121 (atlas) (1937).
- Rightmire, G. P. *The Evolution of Homo erectus* (Cambridge Univ. Press, 1990).
- Wood, B. *Koobi Fora Research Project IV: Hominid Cranial Remains* (Clarendon, Oxford, 1991).
- Tobias, P. V. *Olduvai Gorge IV: The Skulls, Endocranial and Teeth of Homo habilis* (Cambridge Univ. Press, 1991).
- Brown, B. & Walker, A. in *The Nariokotome Homo erectus Skeleton* (eds Walker, A. & Leakey, R.) 161–192 (Harvard Univ. Press, Cambridge, MA, 1993).
- Woo, J. K. *Paleontol. Sinica* **D11**, 1–94 (1962).
- Zhang Yingun *Am. J. phys. Anthropol.* **59**, 21–32 (1982).

- Wang, J., Chen, S., Luo, H. & Zhong, Y. *Acta Anthropol. Sin.* **8**, 172–176 (1989).
- Clochon, R. et al. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Gao, J. *Vertebr. Palasiatica* **13**, 81–88 (1975).
- Leakey, M. D. *Olduvai Gorge III: Excavations in Beds I and II, 1960–1963* (Cambridge Univ. Press, 1971).
- Harris, J. W. K. *Afr. archaeol. Rev.* **1**, 3–31 (1983).
- Clarke, R. J. *World Archaeol.* **21**, 1–12 (1988).
- Swisher, C. C. et al. *Science* **263**, 1118–1121 (1994).
- Wood, B. in *Species, Species Concepts, and Primate Evolution* (eds Kimbel, W. & Martin, L. B.) 485–522 (Plenum, New York, 1993).
- Wood, B. *Nature* **355**, 783–790 (1992).
- Clarke, R. J. *J. hum. Evol.* **19**, 699–736 (1990).
- Howell, F. C. in *Origin of Anatomically Modern Humans* (eds Nitecki, M. H. & Nitecki, D. V.) 253–319 (Plenum, New York, 1994).

ACKNOWLEDGEMENTS. For assistance and permission to conduct research at Longgupo, we thank the Institute of Vertebrate Paleontology and Paleoanthropology (IVPP), the Chinese Academy of Sciences, the National Cultural Relics Bureau, and the Chongqing Museum of Natural History, Sichuan. We specifically acknowledge Li Chuankuei, Jia Lanpo, Zhou Shiwu, J. Olsen and D. Gifford-Gonzalez. Fieldwork was funded by the L.S.B. Leakey Foundation. Laboratory support has been provided by the Wenner-Gren Foundation and the Human Evolution Research Fund at the University of Iowa Foundation. ESR dating was funded by a grant from the NSF. Special thanks to H. Ruiz for composing Figs 1 and 2 and to L. González for helpful discussions.

The neural basis of the central executive system of working memory

Mark D'Esposito*, John A. Detre*, David C. Alsop†, Robert K. Shin*, Scott Atlas† & Murray Grossman*

Departments of * Neurology and † Radiology, University of Pennsylvania Medical Center, Philadelphia, Pennsylvania 19104, USA

WORKING memory refers to a system for temporary storage and manipulation of information in the brain, a function critical for a wide range of cognitive operations. It has been proposed that working memory includes a central executive system (CES) to control attention and information flow to and from verbal and spatial short-term memory buffers¹. Although the prefrontal cortex is activated during both verbal and spatial passive working memory tasks²⁻⁸, the brain regions involved in the CES component of working memory have not been identified. We have used functional magnetic resonance imaging (fMRI) to examine brain activation during the concurrent performance of two tasks, which is expected to engage the CES. Activation of the prefrontal cortex was observed when both tasks are performed together, but not when they are performed separately. These results support the view that the prefrontal cortex is involved in human working memory.

Cerebral activation was measured in six normal, right-handed subjects (3 men, 3 women, aged 22-28 years) with fMRI using blood oxygen level-dependent contrast⁹. The experimental working-memory paradigm consisted of a comparison between dual-task and single-task performance. To test the hypothesis that dorsolateral prefrontal cortex is involved in dual-task performance, we chose two non-working-memory single tasks, a semantic-judgement task and a spatial-rotation task, which predominantly activated posterior brain regions in previous functional neuroimaging studies^{10,11}. The semantic-judgement task required the subject to identify exemplars of a target category ('vegetable') in a series of aurally presented words. The spatial-rotation task required subjects to indicate which of two squares had a dot in the same location, relative to a double line, as a spatially rotated target square¹¹. The dual-task condition required subjects to perform the semantic-judgement task and spatial rotation task concurrently. During the single-task condi-

tion, the mean accuracy of performance on the spatial-location task was 93% and on the semantic-judgement task was 95%. The decrement in performance during the dual-task condition ranged from 0 to 5% for the semantic-judgement task and from 4 to 11% for the spatial task. The protocol for each experimental condition is illustrated in Fig. 1.

When compared to the resting baseline (Fig. 2), the spatial-rotation task produced significantly increased activation in bilateral superior parietal regions (Brodmann's area (BA) 7) and

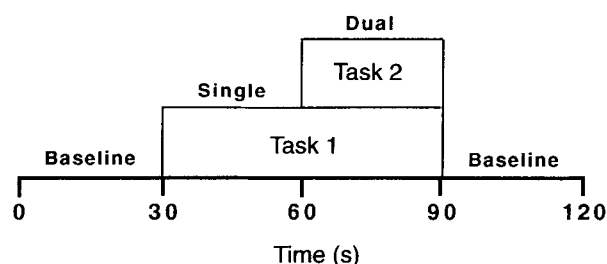


FIG. 1 Illustration of task design. Each activation run began with a 30-s resting baseline, followed by 30 s of either the semantic task or the visuospatial task performed alone (task 1), then 30 s of both tasks performed simultaneously (tasks 1 and 2), and concluded with another 30 s of resting baseline. Each subject performed multiple runs in which the order of task presentation was varied so that task 1 and task 2 alternated between the semantic and spatial task. Subjects were trained on the tasks before the fMRI study. Auditory stimuli were presented over the MRI intercom system at a rate of one word every 2 s. The visual stimuli were presented every 3 s on a rear projection screen by a computer using a LCD panel and an overhead projector. Subjects visualized images on the screen by looking through a mirror attached to the top of the MRI head coil. Subjects responded in the affirmative to each stimulus by dorsiflexion of their toes.

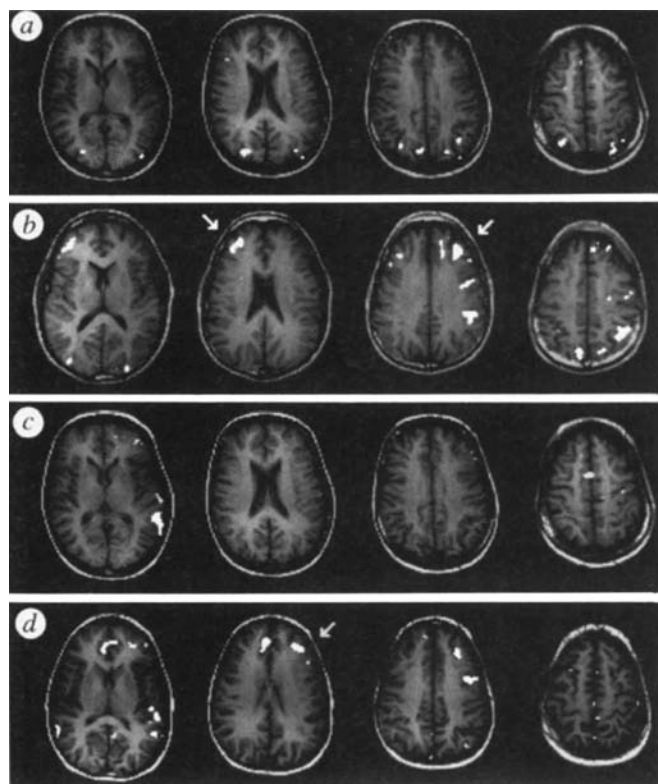


FIG. 2 Regions significantly activated in the single- and dual-task conditions in individual subjects are shown in a-d. Single-task activations are from the same subject; dual-task activations are from two different subjects. Comparison of the single-task versus baseline conditions show activation in the bilateral parieto-occipital regions for the dot-location task (a) and the left temporal region for the semantic-judgement task (c). Regions significantly activated during the dual-task versus single-task conditions are shown in b and d. In the first comparison (dual task minus semantic task), activation is seen in prefrontal cortex, anterior cingulate and premotor cortex (b). In addition, activation is seen in bilateral parieto-occipital regions, which represents activation during the dot-location task. In the second comparison (dual task minus spatial task), activation again is seen in prefrontal cortex and anterior cingulate (d). In addition, activation is seen in bilateral superior temporal gyri, which represents activation during the semantic-judgement task. The fMRI was performed at 1.5 T (G.E. Signa) using a prototype fast gradient system. Four axial slices through dorsolateral prefrontal cortex were chosen from sagittal images. Each activation run consisted of 60 gradient-echo echo-planar images (TR = 2 s, TE = 50 ms, 64 × 64 matrix, 24-cm field of view, 5-mm slices with 5-mm gaps). In-plane motion correction was performed and an upper threshold of 2 s.d. above typical changes in signal intensity with task activation was set to eliminate large signal changes representing residual motion. Activation maps were determined using a pixel-wise t-test for significant signal intensity changes in task versus baseline conditions, including a 4-s delay for vascular transit²⁰. Pixels exceeding a lower threshold for significance were overlaid on axial T1 weighted images; $P < 0.005$, grey; $P < 0.001$, white; negative signal changes not displayed. A 3 × 3 pixel median filter was used to eliminate isolated pixels exceeding these thresholds, resulting in increased significance for the displayed activation. Anatomical location of activated regions was determined using a computer-based stereotaxic coordinate system²¹.

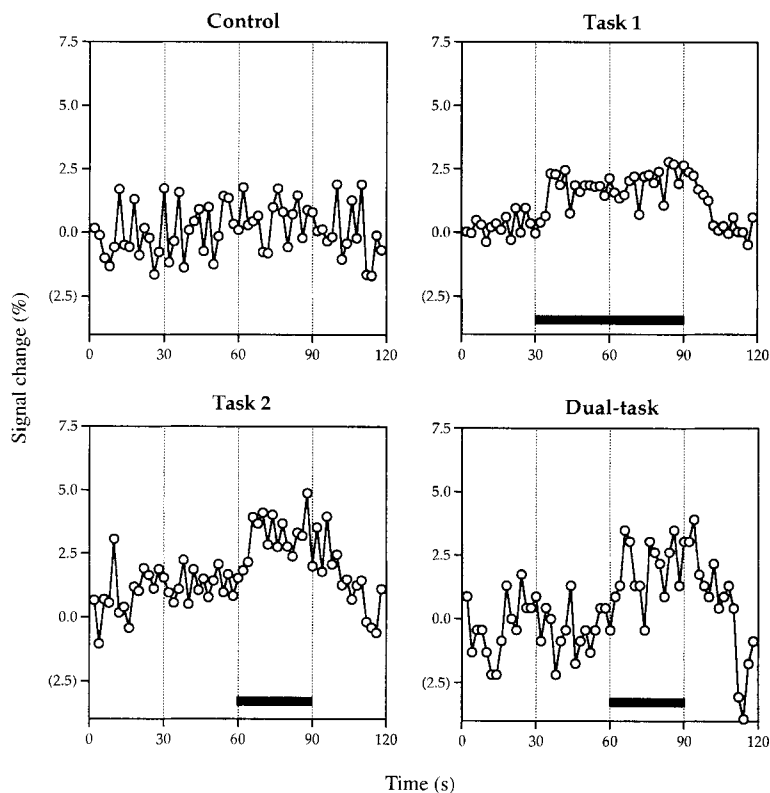


FIG. 3 Plots of the average change in signal intensity over the course of the entire task from selected regions of interest in one subject. 0–30 s, baseline; 30–60 s, single task; 60–90 s dual task; 90–120 s, baseline. Regions of interest were chosen from the parietal lobe (task 1 and task 2) and the dorsolateral prefrontal cortex (dual task), as well as from a control region which was not activated. These plots demonstrate that in task 1, task 2 and the dual task, signal intensity changes correlate with the experimental paradigm illustrated in Fig. 1. No significant task-specific signal-intensity changes occurred in the control region.

lateral occipital regions (BA 19) in all subjects. Significant activity above baseline on the semantic-judgement task (Fig. 2) was limited to the most posterior aspects of superior temporal gyrus (BA 22) in both hemispheres and in the left inferior parietal lobule (BA 39, 40). The full extent of the superior temporal gyrus was not within our chosen slice coverage. We did not observe frontal activation during these single-task conditions.

When the single-task conditions were compared to the dual-task conditions (Fig. 2), all subjects showed significantly increased activation bilaterally in dorsolateral prefrontal cortex (BA 9 and 46). Additional activation was found in the anterior cingulate region (five subjects) and left premotor cortex (two subjects). In these analyses, only one single task could be subtracted from the dual-task condition. Thus posterior activation that was associated with the second single task was also visualized along with frontal activation. Region-of-interest analyses confirmed that signal intensity changes correlated with the tasks (Fig. 3). Individual subject analysis revealed some variation across subjects in the areas of activation within the prefrontal cortex. In four subjects the prefrontal activation was greater in the right hemisphere.

To address the possibility that increased mental effort alone during the dual-task as compared to the single tasks could account for the frontal activation observed, an additional 6 normal, right-handed subjects (3 men, 3 women, age 18–37 years) performed the spatial-rotation task under more difficult conditions by increasing the stimulus frequency. In the first condition, mean accuracy of performance was similar to that during the dual-task paradigm (85%), and in the second condition, there was a significant decrement in performance (74%, $P < 0.005$). Although increasing task difficulty resulted in increased activation in posterior brain regions, no additional prefrontal activation was observed (Fig. 4).

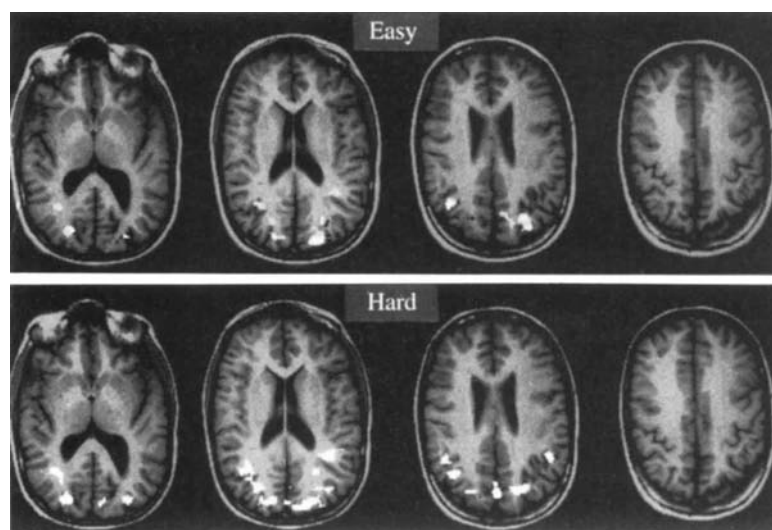
Our results demonstrate recruitment of the dorsolateral prefrontal cortex during the performance of two concurrent tasks, even though neither task activated prefrontal cortex when performed alone. It has been shown that dorsolateral prefrontal regions were among the areas recruited during verbal (that is,

phonological loop) and spatial (that is, visuospatial sketchpad) working-memory challenges in normal humans^{2,6}. Investigations with monkeys have also demonstrated that dorsolateral prefrontal cortex is essential for slave-system working memory^{7,8}. The single tasks used in the present study did not require holding information on-line over a delay period. Although it is conceivable that these (or any) tasks do require working memory to a small extent, it is unlikely that the robust prefrontal activation observed during the dual task is due to a summation of prefrontal activation present, but not detected, in the single tasks. Instead, our findings support the hypothesis that dorsolateral prefrontal cortex is involved in the allocation and coordination of attentional resources, a unique process observed by the CES component of working memory recruited during dual-task performance.

Our findings suggest that there is a partly overlapping neural substrate that underlies both CES and slave-system components of working memory, despite behavioural studies in normal humans that have demonstrated that the CES is a distinct component of working memory, separable from the slave systems¹. Behavioural dissociation of these cognitive processes is also supported by studies of brain-damaged patients who demonstrate significant CES deficits with normally functioning slave systems¹². There are several possible explanations for these findings. First, previous activation studies of slave-system working memory may not have used pure passive memory tasks and may require executive processes^{2,4}. Thus prefrontal activation in these studies may reflect CES functioning. Second, the spatial resolution of fMRI may not be sufficient to identify subregions of the frontal lobe that may subserve different working-memory components. Third, the same neuronal population within prefrontal cortex in different dynamic states may subserve several distinct cognitive operations.

A general problem affecting the interpretation of functional neuroimaging activation studies is that the experimental task is typically more demanding than the baseline or control task. In this study, the dual-task condition was likely to be more demanding than either single-task condition, making it necessary to elim-

FIG. 4 Regions significantly activated in the spatial-rotation task studied with two different levels of a difficulty in a single subject. In these studies, comparison of the same spatial task used in the dual-task studies was made against a sensorimotor control task which required subjects to respond left and right alternatively to three empty squares. The task consisted of four 40-s blocks of the dot-location task alternating with four 40-s blocks of the sensorimotor control task, accumulating 160 gradient-echo echo-planar images in each activation run. Activation maps were obtained using a pixel by pixel correlation analysis to identify pixels whose signal intensity over time followed a predefined task-related reference function²², using a threshold from $r > 0.2$, $P < 0.01$ (grey) to $r > 0.3$, $P < 0.0001$ (white); negative signal changes are not displayed. Both conditions activated identical regions in bilateral parietooccipital regions. There was greater activation (1.3% versus 0.8% signal change, $P < 0.001$) and spatial extent (145 versus 89 pixels, $P < 0.0001$) during the more difficult condition (left) compared to the less difficult task (right). Activation of prefrontal cortex was not observed.



inate the possibility that prefrontal cortex activation was due to a nonspecific increase in mental effort required to perform the dual task, rather than a specific type of working-memory processing, namely, a central executive system. In separate experiments, we addressed this issue by having subjects perform the spatial-rotation task alone at different levels of difficulty. Even at the more difficult condition, when performance was worse than that seen in the dual-task condition, we did not observe any prefrontal activation. We conclude that the prefrontal activation observed is related specifically to dual-task performance.

The anterior cingulate was also activated in most subjects only during the dual-task condition. Anterior cingulate activation has been observed in several cognitive neuroimaging studies^{13, 15}, and this region has been proposed to be part of an anterior attentional system that is critical for response selection among competing, complex contingencies¹⁶. For example, using positron emission tomography¹³, anterior cingulate as well as dorsolateral prefrontal cortex was recruited in a divided-attention task during which subjects had to detect a change in the shape, colour or speed of a visual stimulus. These regions were not recruited when subjects had to attend selectively to only one of these attributes. Simultaneous attention to multiple-stimulus attributes is analogous to performing a dual task and is likely to require the CES of working memory. Thus anterior cingulate cortex activation is also consistent with our hypothesis for the neural basis of the CES. Moreover, recruitment of anterior cingulate and dorsolateral prefrontal cortex, which are anatomically interconnected¹⁷, suggests that the CES may comprise several components. This idea has been proposed in behavioural studies¹⁸, and is supported by single-neuron recording experiments in non-human primates¹⁹.

A potentially important finding in this study is the variability in the distribution of activated regions within the prefrontal cortex. This variability may reflect differing strategies used by neurologically intact subjects to perform a dual task. Investigation of other dual-task paradigms will provide additional information about the adaptive nature of distributed cognitive networks involved in working memory. Another source of variability in activation among subjects may arise from individual differences in the extent to which our dual-task paradigm placed demands on the CES. Studies that carefully monitor individual differences in available resource capacity for dual-task performance are necessary to address this issue. □

4. Cohen, J. et al. *Hum. Brain Mapp.* **1**, 293–304 (1994).
5. McCarthy, G. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 8690–8694 (1994).
6. Swartz, B. et al. *Cereb. Cort.* **3**, 205–214 (1995).
7. Fuster, J. & Alexander, G. *Science* **173**, 652–654 (1971).
8. Goldman-Rakic, P. in *Handbook of Physiology* Vol. 5 (eds Mountcastle, V. B., Plum, F. & Geiger, S. R.) 373–417 (American Physiological Society, Bethesda, 1987).
9. Ogawa, S. et al. *Biophys. J.* **64**, 803–812 (1993).
10. Démonet, J. et al. *Brain* **115**, 1753–1768 (1992).
11. Haxby, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 1621–1625 (1991).
12. D'Esposito, M. et al. *Neuropsychology* (in the press).
13. Corbetta, M., Miezin, F., Dobmeyer, S., Shulman, G. & Petersen, S. *J. Neurosci.* **11**, 2383–2402 (1991).
14. Pardo, J., Pardo, P., Janer, K. & Raichle, M. *Proc. natn. Acad. Sci. U.S.A.* **87**, 256–259 (1990).
15. Raichle, M. et al. *Cereb. Cort.* **4**, 8–26 (1994).
16. Posner, M. & Petersen, S. A. *Rev. Neurosci.* **13**, 25–42 (1990).
17. Goldman-Rakic, P. A. *Rev. Neurosci.* **11**, 137–156 (1988).
18. Van der Linden, M., Coyette, F. & Seron, X. *Cogn. Neuropsychol.* **9**, 301–326 (1992).
19. Goldman-Rakic, P., Chafee, M. & Friedman, H. in *Brain Mechanisms of Perception and Memory: From Neuron to Behavior* (eds Ono, T., Squire, L., Raichle, M., Perret, D. & Fukuda, M.) 445–456 (Oxford Univ. Press, New York, 1993).
20. Kwong, K. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5675–5679 (1992).
21. Talairach, J. & Tournoux, P. *Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme, New York, 1988).
22. Bandettini, P. A., Jesmanowicz, A., Wong, E. C. & Hyde, J. S. *Magn. Reson. Med.* **30**, 161–173 (1993).

ACKNOWLEDGEMENTS. This work was supported by grants from the Charles A. Dana Foundation, the McDonnell-Pew Program in Cognitive Neuroscience, and the US Public Health Service.

Specificity of monosynaptic connections from thalamus to visual cortex

R. Clay Reid & Jose-Manuel Alonso

Laboratory of Neurobiology, The Rockefeller University, New York, New York 10021, USA

IN cortical area 17 of the cat, simple receptive fields are arranged in elongated subregions that respond best to bright (on) or dark (off) oriented contours, whereas the receptive fields of their thalamic inputs have a concentric on and off organization¹. This dramatic transformation suggests that there are specific rules governing the connections made between thalamic and cortical neurons^{1–3} (see ref. 4). Here we report a study of these rules in which we recorded from thalamic (lateral geniculate nucleus; LGN) and cortical neurons simultaneously and related their receptive fields to their connectivity, as measured by cross-correlation analysis^{5,6}. The probability of finding a monosynaptic connection was high when a geniculate receptive field was superimposed anywhere over an elongated simple-cell subregion of the same signature (on or off). However, 'inappropriate' connections from geniculate cells

Received 8 June; accepted 7 September, 1995.

1. Baddley, A. *Working Memory* (Oxford Univ. Press, New York, 1986).
2. Petrides, M., Alivisatos, B., Meyer, E. & Evans, A. *Proc. natn. Acad. Sci. U.S.A.* **90**, 878–882 (1993).
3. Jonides, J. et al. *Nature* **363**, 623–625 (1993).

of the opposite receptive field signature were extremely rare. Together, these findings imply that the outline of the elongated, simple receptive field, and thus of cortical orientation selectivity, is laid down at the level of the first synapse from the thalamic afferents.

Spatiotemporal white-noise stimuli^{7,8} (random checkerboards) were used to make detailed maps of the receptive fields of neurons recorded simultaneously in the lateral geniculate nucleus and the cortex. Figure 1 shows an example from such a dual recording. The geniculate receptive field had a classical on-centre/off-surround organization (Fig. 1a) and the cortical simple receptive field was obliquely oriented, with adjacent on and off subregions (Fig. 1c). To quantify the degree of overlap, we calculated parameters that corresponded to the position and size of the X-cell receptive field⁹ (Fig. 1b) and to the position, orientation, length and width of the simple-cell subregions¹⁰ (Fig. 1d). The on centre of the geniculate neuron (red circle) had a high degree of overlap with the on subregion of the simple cell, as can be seen in both the raw data (Fig. 1c) and the smoothed model receptive field (Fig. 1d).

Independent of the receptive-field measurements, cross-correlation analysis provided a quantitative test for direct monosynaptic connections between simultaneously recorded geniculate and cortical cells. As has been shown in a similar study^{11,12}, positive cross-correlations between LGN and cortical neurons show features that are consistent with monosynaptic connections: fast latencies and rise times, both of around 1–2 ms. The criterion we used for a monosynaptic connection was that the cortical neuron was significantly more likely to fire between 1 and 4.5 ms after a geniculate spike. For example, the neurons whose receptive fields are illustrated in Fig. 1 were strongly correlated at this very fast time scale (Fig. 2a). The fraction of cortical spikes accounted for by this 'monosynaptic' peak^{11,12} (the strength of the correlation) was 7.8% in this example. When the slow events were removed from the cross-correlation (Fig. 2b) the fast monosynaptic peak was found to be highly significant ($P < 10^{-6}$). The statistical test was performed on these 'filtered' data, but the strength of connection was measured by subtracting the baseline from the unfiltered cross-correlations (this method is similar to a shuffle-subtraction¹³).

One potential problem was that cells with the opposite response signature (for example, an off geniculate neuron superimposed with an on simple subregion) were less likely to be co-activated by the visual stimulus. Because the visual stimuli created a negative correlation between the cells' firing on a long

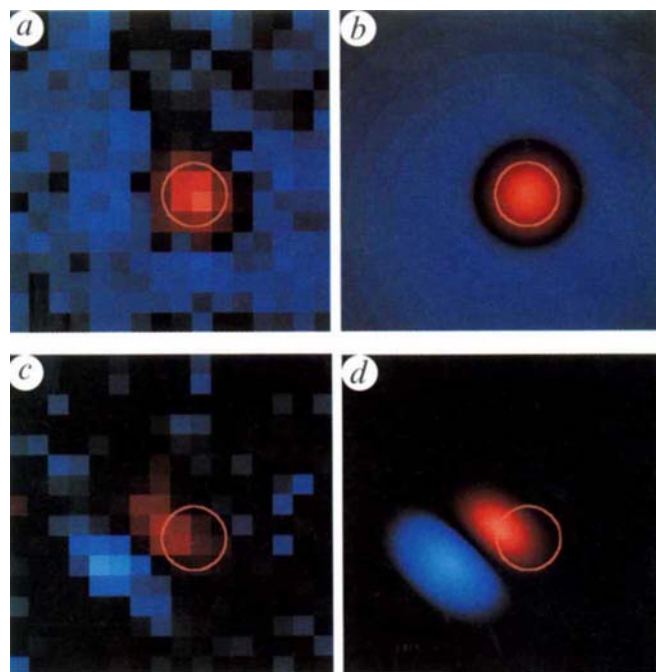


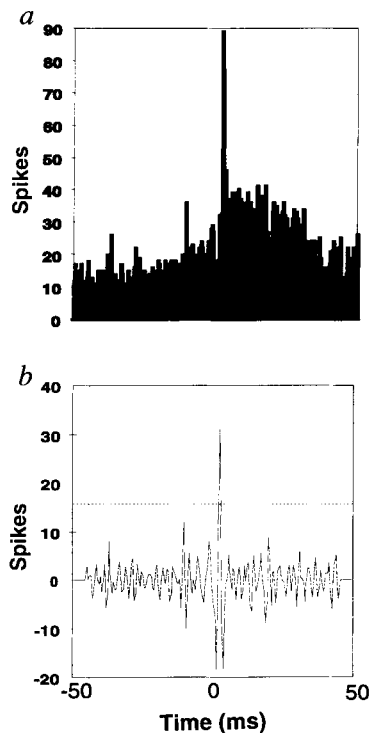
FIG. 1 Receptive field maps of a geniculate X cell (a) and simple cell (c). Hue codes for response sign (red for on responses, blue for off), and brightness codes for response strength. The geniculate cell (a) was on-centre/off-surround: the region in the centre corresponds to the area where the neuron was excited by a light stimulus; the brighter the red, the stronger the excitation. The diffuse blue region (the surround) corresponds to an area where it was weakly excited by a dark stimulus. Black corresponds to areas where the neuron was indifferent to light and dark stimuli. The cortical receptive field (c) has elongated on and off subregions. In b and d, these receptive field maps are fitted to functional forms: a difference of Gaussians⁹ for the geniculate cell, and a Gabor function¹⁰ for the cortical cell. The centre region of the geniculate cell is marked by a red circle with a radius of 1.5 standard deviations of the centre Gaussian. This centre is reproduced on all maps to show the degree of overlap of the two receptive fields.

METHODS. Cats were anaesthetized with ketamine (10 mg per kg, intramuscular) and sodium pentothal (2 mg per kg per h), followed by paralysis with recuronium bromide (0.2 mg h⁻¹, intravenous). Temperature, electrocardiogram, electroencephalogram and expired CO₂ were monitored continuously throughout the experiment. Eyes were dilated with 1% atropine and the nictitating membrane retracted with 10% phenylephrine. The refractive error was measured with a slit ophthalmoscope and corrected with clear contact lenses of the appropriate curvature. Action potentials of single neurons in both the geniculate and area 17 were recorded with plastic-coated tungsten microelectrodes, and were collected on a personal computer running Brainwave software (Broomfield, CO). Receptive-field eccentricities ranged between 5° and 10°. Often, two neurons were recorded simultaneously on either the geniculate or cortical electrode. Visual stimuli were created with a personal computer video-graphics card (Pepper Graphics System, Number Nine Computer Corporation) and presented on a cathode ray tube monitor (NEC Multisync). The white-noise stimulus consisted of a 16 by 16 grid of square pixels. For every frame of the stimulus, each pixel was either black or white according to a binary temporal signal, known as a maximal-length shift register sequence, or m-sequence^{7,14}. The stimulus was updated every two frames of the 100 Hz or 80 Hz video display (every 20 or 25 ms). A complete m-sequence had $2^{15} - 1 = 32,767$ frames and lasted 11 min. The pixel size was 0.4° in most cases, including the data illustrated, and the entire stimulus spanned 6.4°. The receptive-field map is defined as the average stimulus that preceded each action potential as a function of the delay; it is the spatial stimulus that tended to make the neuron fire^{7,8}.

TABLE 1 Connections between overlapping X cells and simple cells

	Connected	Total	Percentage	Strength mean/median (%)
Same-sign	17	27	63%	3.5/2.6
(centred)	6	7	86%	2.1/1.6
Border	5	31	19%	2.4/2.0
Opposite-sign	1	16	6%	0.2/0.2
Total	23	74	31%	3.1/2.4

Summary of connected geniculate/simple-cell pairs as a function of receptive-field overlap. Geniculate X cells were divided into three categories according to the position of their receptive fields relative to the nearest simple cell subregion. If the X-centre was within the central 50% of a simple subregion (along its width) and had the same response signature (on or off) then it was called same-sign. The opposite-sign category was defined similarly, but for the opposite response signature. Otherwise the X cell was said to be along a border of a simple subregion. Most (17 of 27) of the same-sign cells were connected, but only one of 16 opposite-sign cells was. Same-sign geniculate cells were also said to be centred if they were over the strongest simple subregion (as opposed to a weaker flank), within half a standard deviation of the peak along its length, and within the central 40% in width. Of the 7 centred geniculate cells, 6 were connected. The last column shows both the mean and median strengths (see Fig. 2b) of the positive correlations in each category. Because of correlations between thalamic neurons themselves, these values may be overestimates of the connection strengths between thalamic and cortical neurons¹⁵.



time scale (>10 ms), there was occasionally very little baseline on which to observe a 'monosynaptic peak' in the cross-correlation. This problem was less severe with the white-noise input than with drifting gratings, because cells were less dominated by any single feature of the stimulus. Nevertheless, to minimize false negatives we rejected cell pairs whose cross-correlations had very few spikes in the baseline (for both white-noise and grating stimuli, see Fig. 2 legend). This ensured that the two neurons were co-active on a slow, visually evoked time scale frequently enough for a fast monosynaptic influence to be detected.

We recorded from 104 superimposed pairs of geniculate and simple cells for long enough to characterize their receptive fields and to collect cross-correlations with an adequate baseline (Fig. 2a). Here we consider those 74 pairs for which the geniculate cells were X cells with centres similar in size to the cortical subfields. Of these 74 X cells, 23 were positively correlated with a superimposed simple cell ($P < 0.016$; see Fig. 2b). The strength of the positive correlations ranged from 0.2% to 10.3%, with an average of 3.1%. For the remaining 30 pairs the geniculate cells had larger receptive fields and were probably Y cells. Eight of these 30 cells were positively correlated with a simple cell^{11,12}, but there was no clear relation between receptive-field overlap and the probability of connection. These cells will not be considered further in this study.

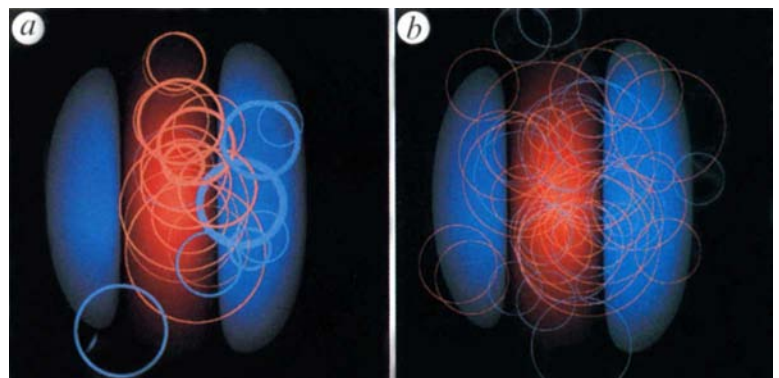
FIG. 2 Cross-correlation between the geniculate and cortical neurons illustrated in Fig. 1. The ordinate gives the number of times the cortical neuron fired as a function of the delay between geniculate and cortical firing (+50 ms, in steps of 0.5 ms). The cortical cell was most likely to fire between 1 and 4.5 ms after a geniculate action potential. The strength of a connection or the contribution^{10,11}—the percentage of the cortical spikes that were accounted for by the 'monosynaptic' peak (1–4.5 ms) above the baseline—was 7.8%. Data were collected over 11 min during white-noise stimulation. Geniculate spikes, 9,740; cortical spikes, 2,176. The baseline was defined as the total number of spikes between -4.5 and -1.0 ms. To minimize false negatives, only cross-correlations that had a baseline of >50 spikes were included. This value was chosen because, of the 15 cross-correlations that had <50 spikes (not included in our sample), only one had a statistically significant peak, whereas 5 of the 12 cross-correlations with baselines between 50 and 100 had significant peaks. Cross-correlations were calculated from periods of visual stimulation either with white noise (12 positive out of 37), or with oriented gratings (11 positive out of 37). For the grating data, the strengths calculated from the shuffle-subtractions¹³ (average 2.0, median 1.4) and with the subtraction of the baseline (average 1.6, median 1.8) were very similar. b, Test of statistical significance. The same cross-correlation was bandpass filtered between 75 and 700 Hz to capture only the fastest, 'monosynaptic' correlations. These frequencies are faster than most visual responses, so this procedure removed stimulus-dependent correlations just as effectively as a shuffle-subtraction. The broken line indicates the significance level set at a probability of 0.2%, assuming a normal distribution. As there are eight samples in the range 1–4.5 ms, this corresponds to a probability of $P < 1.6\%$. This criterion was used for all of the cross-correlations in the study.

The mapping techniques used here allowed us to extend previous results^{11,12} by relating the receptive-field positions of the geniculate and cortical neurons to the probability of finding a connection between them. Ideally, we would record from all the geniculate cells whose receptive fields overlap a single simple cell. Instead, data collected from different simple-cell/X-cell pairs were compared by using their functional forms (Fig. 1b, d), which were rotated and stretched onto an idealized receptive field (Fig. 3). If the strongest simple subregion was off, the signs of the cortical and geniculate receptive fields were inverted. The 74 X cells were densely scattered over the model simple receptive field. In Fig. 3 we show just the 23 X cells that were positively correlated. The size of the circles represents the X-cell centre, the colour represents the signature relative to the central, strongest cortical subregion (red, same sign; blue, opposite sign), and the thickness represents the strength of the correlation.

The centres of most of the connected geniculate neurons were well superimposed with a simple cell subregion of the same sign. This demonstrates that individual simple cells receive input from both on and off geniculate cells whose centres are aligned along the axis of each subregion.

To tabulate the probability of finding a monosynaptic connection, cell pairs were grouped according to the spatial overlap between the geniculate centre and the nearest cortical subregion. The percentage of positively correlated pairs was calculated for

FIG. 3 Summary of the spatial relation between simple and geniculate receptive fields of the 23 functionally connected cell pairs (those with positive correlations). The receptive field fitted for each simple cell was transformed (rotated and stretched) into an idealized receptive field. The geniculate centres are shown in relation to the idealized simple cell, which has a strong subregion in the centre (red) and two flanks of unequal strength (blue). Geniculate centres with the same sign as the strongest simple subregion are shown in red, those with the same sign as the flanks in blue. The diameter indicates the relative receptive-field size (1.5 standard deviations), and thickness indicates the strength of the correlation, from 0.2% to 10.3%. b, Spatial relation between simple and geniculate receptive fields of the 51 pairs that were not functionally connected.



three groups: opposite-sign, border and same-sign (Table 1). Most (17 of 27) of the same-sign geniculate cells were connected, as would be expected according to the original Hubel and Wiesel model¹. A small fraction of the border cells (5 of 31) had positive correlations, but only one of the opposite-sign geniculate cells in our sample was connected. The strength of this 'inappropriate' connection was the weakest seen, only 0.2%. There was no other association between receptive-field overlap (for example, same-sign versus border) and the strength of the positive correlations.

Because a significant fraction of the same-sign cells were not connected, we examined several subcategories to see if any had a higher probability of being connected. There was no asymmetry between on and off; in particular, off-centre X cells were connected to simple cells whose strongest subregion was off and to simple cells with off flanks. An almost equal percentage of the same-sign cells that were over the strongest subregion were connected (12 of 20) compared to those over a flank (5 of 7). However, of the same-sign X cells that were over the strongest subregion, those that were very well centred in length and width had the highest proportion of positive correlations (6 of 7).

In summary, the connections made between geniculate X cells and simple cells were very specific. If the geniculate receptive field centre overlapped an appropriately signed simple subregion, the probability of connection was high (63%). The probability was highest if the X cell was well centred over the strongest simple subregion (86%), and decreased in both width and length. There were very few exceptions to these simple rules: only one of the 23 connected X cells was superimposed with a subregion

of the wrong response sign. Thus both the subfield organization and the orientation of simple receptive fields are well established by the monosynaptic thalamic inputs. These findings imply specificity at the finest level, not only in terms of the afferents that project to a given column², but also in terms of the selective connections made with different cortical neurons within a column. The specificity of the thalamocortical projection does not deny a role for intracortical connections, both excitatory and inhibitory, in strengthening orientation selectivity. It does, however, demonstrate the importance of the geniculate input in the creation of simple receptive fields. □

Received 16 June; accepted 3 October 1995.

- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160**, 106–154 (1962).
- Chapman, B., Zahs, K. R. & Stryker, M. P. *J. Neurosci.* **11**, 1347–1358 (1991).
- Ferster, D. *J. Neurosci.* **7**, 1780–1791 (1987).
- Sillito, A. M. *J. Physiol., Lond.* **250**, 305–329 (1975).
- Alonso, J. M., Reid, R. C. & Wiesel, T. N. *Soc. Neurosci. Abstr.* **19**, 425 (1993).
- Reid, R. C., Alonso, J. M. & Wiesel, T. N. *Soc. Neurosci. Abstr.* **20**, 425 (1994).
- Reid, R. C. & Shapley, R. M. *Nature* **356**, 716–718 (1992).
- Jones, J. P. & Palmer, L. A. *J. Neurophysiol.* **58**, 1187–1211 (1987).
- Rodieck, R. W. *Vision Res.* **5**, 583–601 (1965).
- Jones, J. P. & Palmer, L. A. *J. Neurophysiol.* **58**, 1233–1258 (1987).
- Tanaka, K. *J. Neurophysiol.* **49**, 1303–1318 (1983).
- Tanaka, K. *Vision Res.* **25**, 357–364 (1985).
- Perkel, D. H., Gerstein, G. L. & Moore, G. P. *Biophys. J.* **7**, 419–440 (1967).
- Sutter, E. *Adv. Meth. Physiol. Systems Model.* **1**, 303–315 (Univ. Southern California, 1987).
- Usrey, W. M., Alonso, J. M. & Reid, R. C. *Soc. Neurosci. Abstr.* **21**, 21 (1995).

ACKNOWLEDGEMENTS. Both authors contributed equally to this work. We thank T. N. Wiesel for discussion, J. Hirsch for comments on the manuscript, W. M. Usrey for help with some of the experiments, and C. Gallagher and K. McGowan for technical assistance. This study was supported by the NIH, the Klingenstein Fund, Fulbright/MEC, and the Charles H. Revson Foundation.

Displacement of corticotropin releasing factor from its binding protein as a possible treatment for Alzheimer's disease

Dominic P. Behan*, Stephen C. Heinrichs*, Juan C. Troncoso^{†,‡}, Xin-Jun Liu*, Claudia H. Kawas[‡], Nicholas Ling* & Errol B. De Souza*[§]

* Neurocrine Biosciences, Inc., 3050 Science Park Road, San Diego, California 92121, USA

Neuropathology Laboratory, Departments of † Pathology and

‡ Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

In Alzheimer's disease (AD) there are dramatic reductions in the content of corticotropin releasing factor (CRF)^{1–4}, reciprocal increases in CRF receptors^{1,2}, and morphological abnormalities in CRF neurons⁵ in affected brain areas. Cognitive impairment in AD patients is associated with a lower cerebrospinal fluid concentration of CRF⁶, which is known to induce increases in learning and memory in rodents^{7–9}. This suggests that CRF deficits contribute to cognitive impairment. The identification in post-mortem brain of CRF-binding protein (CRF-BP)^{10,11}, a high-affinity binding protein that inactivates CRF, and the differential distribution of CRF-BP¹² and CRF receptors¹³, provides the potential for improving learning and memory without stress effects of CRF receptor agonists¹⁴. Here we show that ligands that dissociate CRF from CRF-BP increase brain levels of 'free CRF' in AD to control levels and show cognition-enhancing properties in models of learning and memory in animals without the characteristic stress effects of CRF receptor agonists.

Previous studies have defined the characteristics and distribution of CRF-BP in rat brain¹², sheep brain¹⁰ and human plasma^{15–17}. Our initial studies focused on the identification and characterization of CRF-BP in post-mortem human brain. Brains were obtained post-mortem from ten demented individuals (NINCDS clinical diagnosis criteria¹⁸) with pathologically confirmed AD (NINCDS pathological diagnosis criteria¹⁹) (7 males and 3 females; age 74 ± 3.5 years (range 59–89 years); post-mortem delay 6.9 ± 0.7 h (range 3–12 h) and ten age-matched neurologically normal controls (2 males and 8 females; age 76 ± 3.1 years (range 52–85 years); post-mortem delay 6.9 ± 0.6 h (range 5–10 h)). All AD cases showed abundant amyloid deposits, neuritic plaques and variable densities of neurofibrillary tangles in neocortex, fulfilling NINCDS criteria¹⁹. The control cases showed no evidence of neocortical pathology. Most of the CRF-BP-like immunoreactivity (CRF-BP-IR) (>85%) in human cerebral cortex was membrane associated as a result of the presence of activity in the membrane pellets after centrifugation of the tissue extracts. These data are in accordance with studies in sheep-brain homogenates¹⁰ and in slide-mounted sections of rat brain¹², demonstrating a predominantly membrane/neuronal localization of the CRF-BP. The membrane-associated form of the CRF-BP in human brain has pharmacological characteristics similar to those of the recombinant form of the soluble plasma CRF-BP (Table 1).

The levels of CRF and CRF-BP in the cerebral cortex of Alzheimer's patients and age-matched controls were then compared. In agreement with previous observations^{1–4}, there were dramatic reductions in the concentration of CRF-IR in the frontal, parietal and temporal cerebral cortex, with a tendency for a non-significant decrease in occipital cortex (Fig. 1a). Conversely, there were no significant differences seen in CRF-BP-IR in any of the four cerebrocortical areas examined in AD, regardless of whether the CRF-BP was assayed using a ligand immunoradiometric assay or enzyme-linked immunosorbent assay (ELISA) (Fig. 1b). Furthermore, the pharmacological profile of the CRF-BP in AD tissue was similar to that previously described for age-matched controls and the recombinant form of CRF-BP (Table 1). These data provide evidence for the

§ To whom correspondence should be addressed.

TABLE 1 Affinities of CRF and related peptides at the human CRF-BP and CRF-R

Peptide	50% Inhibitory concentration (IC ₅₀) (nM)				Activity in behavioural models	
	CRF-BP in cerebral cortex		Recombinant human CRF-BP	Recombinant human CRF-R		
	Age-matched controls	Alzheimer's Disease			Cognitive enhancer	Anxiogenic effects
Human CRF	0.30	0.18	0.19	1.0	Yes	Yes
α-Helical CRF (9–41)	0.16	0.04	0.20	9.9	ND	ND
Human CRF (6–33)	1.1	0.91	0.89	>10,000	Yes	No
Human CRF (9–33)	3.4	3.2	1.2	>10,000	Yes	No
Ovine CRF	667	595	471	2.4	ND	ND
Ovine CRF (9–33)	>10,000	>10,000	>10,000	>10,000	No	No

CRF-BP was measured by a two-site ligand immunoradiometric assay (LIRMA)¹⁰. The affinity of various peptides for the cloned human pituitary CRF receptor (CRF-R)²⁵ was determined in membrane preparations of stable transfectants of the receptor in Ltk⁻ mouse fibroblast cells using a previously characterized radioligand binding assay²⁶. The learning and memory effects were evaluated in the Morris watermaze²¹ and Y-maze visual discrimination⁷ tests, and the effects on anxiety were evaluated in the elevated plus maze test²². Peptides not evaluated in the behavioural assays are indicated by ND.

presence of CRF-BP in normal and AD brain tissue, and show that deficits in CRF seen in AD are not accompanied by a decrease in CRF-BP. This may be due to decreased CRF synthesis and/or increased degradation rather than loss of CRF neurons. Alternatively, if CRF neuronal loss does occur in AD then the data would suggest that CRF-BP could be localized to non-CRF neurons or non-neuronal cells. In support of this, a localization of CRF-BP to non-CRF cells has been reported in rat brain¹². In both control and AD groups, neither age of the patient nor post mortem delay correlated significantly with CRF-IR or CRF-BP-IR levels. Similarly, no correlation was observed between gender of the patient and CRF-IR or CRF-BP-IR levels in any of the cerebrocortical areas examined, with the exception of CRF-IR measurements in the parietal cortex which showed a significant difference between males and females.

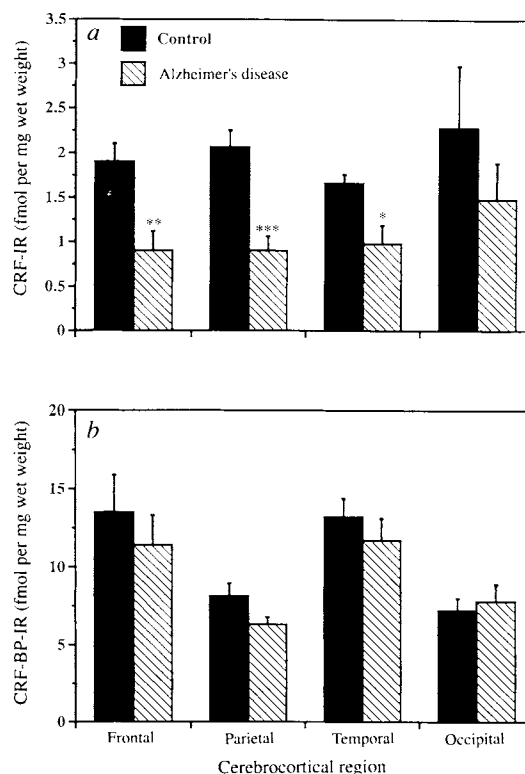
To determine the nature of the interaction of CRF with CRF-BP in human brain tissue, the proportion of CRF complexed to the CRF-BP and the 'free' pool of CRF were measured. Assays specific for 'total CRF' and CRF bound to the binding protein

demonstrated that ~40% and 60% of the 'total CRF' was complexed to the CRF-BP in normal and AD cerebrocortical extracts, respectively (Fig. 2). In both normal and AD cerebral cortex, endogenous CRF was bound to the CRF-BP in a reversible manner, as treatment of the tissue with high-affinity CRF-BP ligands, such as α -helical CRF (9-41) or human CRF (6-33)²⁰, displaced CRF from the CRF-BP as detected by decreases in the CRF/CRF-BP complex in the tissue and increases in the 'free' pool of endogenous CRF in the supernatants. These *in vitro* data provide strong evidence that the CRF-BP ligands used here displace CRF from the CRF/CRF-BP complex and increase the 'free' concentrations of CRF. Treatment of the AD cerebral cortex with a CRF-BP ligand replenished the 'free' concentrations of CRF to normal levels seen in age-matched controls (Fig. 2).

The beneficial cognition-enhancing effects and potential anxiety-related side effects of activation of CRF-sensitive neurons were assessed in rats. Daily pre-test administration of a high-affinity CRF receptor agonist, such as human CRF itself, by

FIG. 1 a, CRF-IR and b, CRF-BP-IR in discrete regions of cerebral cortex of AD patients and controls. All values are means $\pm$ s.e.m. (N=9 or 10 subjects per group). Data are presented as fmol per mg wet weight of tissue. Data were analysed for differences using Student's t-test. Comparable, statistically significant reductions in CRF-IR and no significant alterations in CRF-BP-IR were obtained when the data were calculated as CRF-IR or CRF-BP-IR levels per mg protein. Significant differences from control group are indicated at * P <0.05, ** P <0.015 and *** P <0.005.

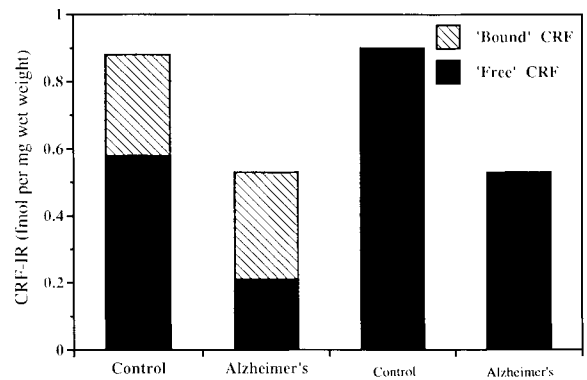
METHODS. Samples of cortical grey matter from standardized regions of frontal, parietal, temporal and occipital lobes were subdivided from brain tissues frozen at autopsy and stored at -70°C . Brain samples were weighed and homogenized in 5 ml of 10% sucrose. For each sample, 1 ml was centrifuged at 10,000 g for 10 min at room temperature and the resultant membrane pellets were reconstituted in SPEA (50 mM sodium phosphate, 0.1 M NaCl, 25 mM EDTA, 0.1% sodium azide buffer, pH 7.4, containing 0.25% bovine serum albumin (BSA) and 1% of the membrane-solubilizing detergent NP-40) for subsequent analysis of CRF-BP-IR using a previously characterized ligand immunoradiometric assay (LIRMA)¹⁰ or a two-site ELISA as described below. CRF was measured by radioimmunoassay as previously described². ELISA plates were coated with 10 $\mu\text{g ml}^{-1}$ of human CRF-BP monoclonal antibody (100 μl) in 50 mM sodium bicarbonate buffer, pH 9.5, for 2 h at 37 $^{\circ}\text{C}$. Plates were then washed once with 100 μl per well of TTBS (20 mM Tris-HCl, 150 mM NaCl, 0.05% Tween-20, pH 7.5) and subsequently blocked with 1% casein TBS (20 mM Tris-HCl, 150 mM NaCl, pH 7.5) for 1 h at room temperature. Samples and standards (100 μl) were diluted in TTBS/1% BSA and then exposed to the CRF-BP monoclonal antibody for 1 h at room temperature. Plates were then washed five times with TTBS and then exposed to a second polyclonal rabbit anti-human CRF-BP antibody (code 5144 diluted 1:1,000 in TTBS/1% BSA) for 1 h at room temperature. Plates were again washed five times with TTBS and exposed to an affinity-purified goat anti-rabbit (GAR) horseradish peroxidase linked second antibody for 1 h at room temperature. Plates were finally washed again five times with TTBS and then developed for 5 min by the addition of 100 μl of TMB microwell



peroxidase substrate solution (Kiegegaard and Perry). The reaction was stopped by the addition of 100 μl of 0.2 M sulphuric acid, and the colour was read in a BioRad plate reader at 450 nm.

FIG. 2 Effect of the CRF-BP ligand on the relative concentrations of 'free CRF' and CRF bound to the CRF-BP in the cerebral cortex of AD patients and age-matched controls. Data represent triplicate determinations from a pool of five cerebrocortical samples each from AD patients and controls.

METHODS. CRF was measured by ELISA as follows. Briefly, ELISA plates were coated with protein G purified sheep anti-CRF antibody ($20 \mu\text{g ml}^{-1}$ for 2 h at 37°C) diluted in 50 mM sodium bicarbonate buffer, pH 9.5. Plates were washed once with TTBS and then blocked with 1% casein TBS for 1 h at room temperature. Sample ($100 \mu\text{l}$) (extracted as described in Fig. 1) and standard were then added to each well and allowed to bind at room temperature. The remainder of the assay was performed as described for the CRF-BP ELISA. 'Bound CRF' was estimated after capture of the CRF/CRF-BP complex in wells that had been precoated with anti-human CRF-BP monoclonal antibody ($5 \mu\text{g ml}^{-1}$ in 50 mM sodium bicarbonate buffer, pH 9.5) followed by detection of the bound CRF with RC-70 anti-human CRF antibody, essentially as described for the 'total CRF' ELISA. The effect of a high-affinity CRF-BP ligand to displace 'bound CRF' was assessed after monoclonal capture of the CRF/CRF-BP complex in the presence of



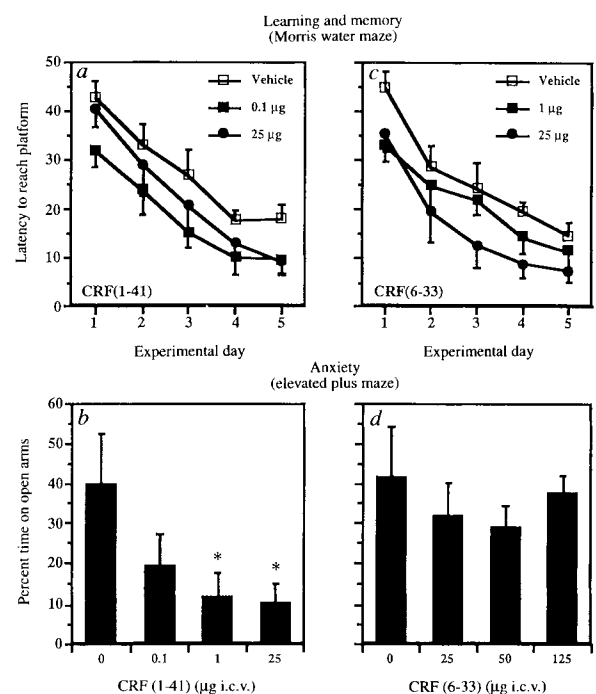
50 nM α -helical CRF (9–41). Displaced 'free CRF' was measured in the supernatants remaining after CRF-BP monoclonal antibody capture by CRF ELISA.

an intracerebroventricular (i.c.v.) route significantly enhanced learning in the Morris watermaze test of spatial learning and memory²¹ at all the doses tested (0.1 – $25.0 \mu\text{g}$) with optimal effects at the $0.1 \mu\text{g}$ dose (Fig. 3a). These data are consistent with previous reports of cognition-enhancing effects of CRF to improve retention of shock-motivated avoidance learning tasks^{8,9} and an appetitively motivated visual discrimination task⁷. However, the same doses of CRF that produced increases in learning and memory also produced anxiety, as shown by dose-dependent suppression of exploration of the open arms in the elevated plus maze test^{14,22} (Fig. 3b). These same doses of CRF are also associated with increases in locomotor activity and generalized arousal and emotionality¹⁴. Thus the benefits of using a CRF-receptor agonist for the treatment of the cognitive deficits seen in AD are limited owing to its associated side effects and low therapeutic index.

Central administration of a selective CRF-BP ligand, human CRF (6–33)²⁰, at an optimal dose of $25 \mu\text{g}$ of the peptide signifi-

cantly increased spatial learning in the Morris watermaze test as measured by both the speed to locate a submerged target (Fig. 3c) and duration of searching the target quadrant. In the appetitively motivated, visual discrimination task, which is known to be sensitive to the performance-enhancing properties of human CRF⁷, administration of human CRF (6–33) to matched groups of subjects exhibiting chance performance on the initial day-1 trials significantly increased the rate of correct responses during the first two days of training ($71 \pm 4\%$ at $5 \mu\text{g}$, and $72 \pm 7\%$ at $25 \mu\text{g}$) relative to vehicle-treated controls ($53 \pm 6\%$), and were not significantly different from controls on subsequent days once vehicle-treated animals achieved a high level of performance (60–70% correct responses). Human CRF (9–33), which also has high affinity for the CRF-BP²⁰ and is devoid of any intrinsic activity at the CRF receptor, produced significant memory-enhancing effects in both models at doses comparable to those of human CRF (6–33), but the inactive CRF-BP/CRF receptor ligand ovine CRF (9–33)²⁰ had no effect on the learning and

FIG. 3 Comparison of the learning and memory and anxiety effects of a CRF-receptor agonist and a CRF-binding protein (CRF-BP) ligand. Mean $\pm$ s.e.m. latency to reach a submerged platform in the Morris watermaze test of spatial learning²¹ (a, c) or percent time spent exploring open arms of the elevated plus maze test of anxiety²² (b, d) in rats treated 15 min before testing with either the CRF receptor agonist (i.c.v. injection of human CRF at 0, 0.1, 1, or $25 \mu\text{g}$) (a, b) or the CRF-BP ligand (i.c.v. injection of human CRF (6–33) at 0, 1, 5, 25, 50 or $125 \mu\text{g}$) (c, d) ($N=7$ – 10 rats per group). Testing in the Morris maze (120 cm diameter, 45 cm high, filled with water at 23°C to a depth of 15 cm) consisted of four consecutive trials per day with a maximum trial duration of 60 s. All treatment groups exhibited matched performance on trial 1 of the first day (latency to reach platform: vehicle, 55 ± 3 s; $25 \mu\text{g}$ human CRF (6–33), 56 ± 2 s) such that group differences arising on either the first day of acquisition (vehicle, 38 ± 4 s; $25 \mu\text{g}$ human CRF (6–33), 29 ± 1 s) or on subsequent days indicate improvement in performance. Data from the Morris maze test were evaluated for significant differences between treatment groups using two-way analysis of variance (ANOVA) and demonstrated significant improvement in the performance following treatment with either human CRF ($F(1,64)=4.74$, $P<0.05$) or human CRF(6–33) ($F(2,108)=3.26$, $P<0.05$); there was a significant improvement in performance over time ($F(4,64)=26.67$, $F(4,108)=44.93$, both $P<0.0001$) but no significant interaction between drug treatment and time for either treatment group. In a replication of the watermaze experiment (c), a separate group of subjects ($N=16$) given either vehicle or a $25 \mu\text{g}$ dose of human CRF (6–33) over five training days were tested in a single 120-s 'probe' trial on day 6 without injections and with the submerged platform removed from the pool. The observed increase in duration of target quadrant searching on the part of peptide-treated subjects relative to controls (vehicle, $37 \pm 3\%$ of total time; $25 \mu\text{g}$ human CRF (6–33), $47 \pm 3\%$, $P<0.03$) indicates unambiguous enhancement of spatial learning. Testing in the elevated plus maze (computer-automated apparatus elevated 40 cm above the floor with arms of 10×50 cm which are



either enclosed within 40-cm walls or left open) lasted for 5 min using maze-naïve subjects. Data from the elevated plus maze test were analysed for significant differences using Student's *t*-test; significant differences from vehicle-treated group are indicated, * $P<0.05$.

memory models (Table 1). Note that, in marked contrast to human CRF, memory-enhancing doses of the CRF-BP ligand human CRF (6–33) (Fig. 3d) and human CRF (9–33), as well as doses two- to fivefold higher (50–125 µg), did not significantly alter performance in the elevated plus maze test of anxiety or produce any overt behavioural alterations similar to those seen with a CRF-receptor agonist (data not shown). Furthermore, the cognition-enhancing properties of the CRF-BP ligand inhibitors do not appear to be due to non-selective effects of the compounds on performance, because efficacious doses of human CRF (6–33) did not alter swimming speed as measured in the Morris maze, or overall activity as measured by cumulative arm entries in the elevated plus maze test (data not shown). Overall, these data demonstrate a clear-cut functional dissociation of the efficacious cognition-enhancing effects from anxiogenic/arousal side effects for a CRF-BP ligand underscoring the therapeutic potential of this mechanism for the treatment of cognitive deficits. This functional dissociation for the CRF-BP ligand inhibitors is more likely to be due to the relative enrichment of the CRF-BP in brain areas involved in learning and memory processes (such as the cerebral cortex and hippocampus) than in some limbic and brainstem areas involved in emotional responses. In contrast, the widespread localization of CRF receptors in brain nuclei subserving both functions would explain the overlap between doses of CRF-increasing learning and memory and those producing anxiety.

In summary, we have identified CRF-BP in human brain and demonstrated that a large proportion of CRF in normal human brain might be in the form of complexes to the CRF-BP and so be biologically inactive. Furthermore, because brain CRF-BP remains unaltered despite diminished CRF content in AD brains, the levels of 'free CRF' in AD are even lower. The behavioural studies demonstrating cognition-enhancing effects of selective CRF-BP ligands in models of learning and memory in animals, without the production of any overt anxiogenic actions or arousal, suggest that this mechanism is useful for the develop-

ment of therapeutic agents for the symptomatic treatment of memory deficits seen in AD. Furthermore, selective ligands that dissociate CRF from the CRF/CRF-BP complex, and thereby increase synaptic levels of 'free CRF', may have wide applicability for the treatment of disorders and conditions associated with low levels of CRF in the brain^{23,24}. □

Received 10 March; accepted 14 September 1995.

1. De Souza, E. B. *Hosp. Pract.* **23**, 59–71 (1988).
2. De Souza, E. B. *et al. Nature* **319**, 591–595 (1986).
3. Bissette, G. *et al. J. Am. med. Ass.* **254**, 3067–3069 (1985).
4. Whitehouse, P. J. *et al. Neurology* **37**, 905–909 (1987).
5. Powers, R. E. *et al. Synapse* **1**, 405–410 (1987).
6. Pomara, N. *et al. Biol. Psychiat.* **26**, 500–504 (1989).
7. Koob, G. F. & Bloom, F. E. *Fedn Proc.* **44**, 259–262 (1985).
8. Liang, K. C. & Lee, E. H. Y. *Psychopharmacology* **96**, 232–236 (1988).
9. Diamant, M. & de Wied, D. *Neuroendocrinology* **57**, 1071–1081 (1993).
10. Behan, D. P. *et al. Ann. N. Y. Acad. Sci.* **697**, 1–8 (1993).
11. Potter, E. *et al. Nature* **349**, 423–426 (1991).
12. Potter, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 4192–4196 (1992).
13. De Souza, E. B. *et al. J. Neurosci.* **5**, 3189–3203 (1985).
14. Koob, G. F. *et al. in Corticotropin-Releasing Factor* (eds Chadwick, D. J., Marsh, J. & Ackrill, K.) 277–295 (Wiley, Chichester, UK, 1993).
15. Linton, E. A., Wolfe, C. D. A., Behan, D. P. & Lowry, P. J. *Clin. Endocr.* **28**, 315–324 (1988).
16. Orth, D. N. & Mount, C. D. *Biochem. biophys. Res. Commun.* **143**, 411–417 (1987).
17. Suda, T. *et al. J. clin. Endocr. Metab.* **67**, 1278–1283 (1988).
18. McKhann, G. *et al. Neurology* **34**, 939–944 (1984).
19. Khachaturian, Z. S. *Archs Neurol.* **42**, 1097–1105 (1985).
20. Sutton, S. W. *et al. Endocrinology* **136**, 1097–1102 (1995).
21. Stewart, C. A. & Morris, R. G. M. *in Behavioral Neuroscience* (ed. Saghal, A.) 107–122 (IRL, Oxford Univ. Press, 1993).
22. Pellow, S. *et al. J. Neurosci. Meth.* **14**, 149–167 (1985).
23. De Souza, E. B. *A. Rep. med. Chem.* **25**, 215–224 (1990).
24. Chrousos, G. P. & Gold, P. W. *J. Am. med. Ass.* **267**, 1244–1252 (1992).
25. Chen, R., Lewis, K. A., Perrin, M. H. & Vale, W. W. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8967–8971 (1993).
26. De Souza, E. B. *J. Neurosci.* **7**, 88–100 (1987).

ACKNOWLEDGEMENTS. D.P.B. and S.C.H. contributed equally to this study. We thank D. E. Grigoriadis for data on the pharmacological profile of CRF-related peptides at the human CRF receptor; W. Vale for antibodies for the measurement of CRF and CRF-binding protein; J. Rivier for ovine CRF (9–33); F. Bloom, G. Koob, W. Vale and S. Watson for critical reading of the manuscript and suggestions; and D. Wickham for technical assistance. This work was supported in part by a National Institute of Neurological Disorders and Stroke Small Business Innovative Research (SBIR) grant. Clinical and neuropathological assessments of AD cases were supported by grants from the NIH and the Charles A. Dana Foundation. J.C.T. is the recipient of a grant from the Alzheimer's Association.

Urocortin, a mammalian neuropeptide related to fish urotensin I and to corticotropin-releasing factor

Joan Vaughan, Cynthia Donaldson, Jackson Bittencourt*, Marilyn H. Perrin, Kathy Lewis, Steven Sutton, Raymond Chan*, Andrew V. Turnbull, David Lovejoy, Catherine Rivier, Jean Rivier, Paul E. Sawchenko* & Wylie Vale†

The Clayton Foundation Laboratories for Peptide Biology,

* Laboratory of Neuronal Structure and Function, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037, USA

CORTICOTROPIN-RELEASING factor (CRF), a peptide first isolated from mammalian brain¹, is critical in the regulation of the pituitary–adrenal axis, and in complementary stress-related endocrine, autonomic and behavioural responses². Fish urotensin I and amphibian sauvagine were considered to be homologues³ of CRF until peptides even more closely related to CRF were identified in these same vertebrate classes^{4,5}. We have characterized another mammalian member of the CRF family and have localized its urotensin-like immunoreactivity to, and cloned related complementary DNAs from, a discrete rat midbrain region. The deduced

protein encodes a peptide that we name urocortin, which is related to urotensin (63% sequence identity) and CRF (45% sequence identity). Synthetic urocortin evokes secretion of adrenocorticotrophic hormone (ACTH) both *in vitro* and *in vivo* and binds and activates transfected type-1 CRF receptors^{6–9}, the subtype expressed by pituitary corticotropes. The coincidence of urotensin-like immunoreactivity with type-2 CRF receptors^{10–13} in brain, and our observation that urocortin is more potent than CRF at binding and activating type-2 CRF receptors, as well as at inducing c-Fos (an index of cellular activation) in regions enriched in type-2 CRF receptors, indicate that this new peptide could be an endogenous ligand for type-2 CRF receptors.

To identify the principal sites of expression of molecules related to urotensin I, we used specific polyclonal antisera and screened rat tissues by radioimmunoassay and immunohistochemistry. In the adult rat brain, the dominant cellular localization of urotensin-like immunoreactivity is in the Edinger–Westphal nucleus (EW in Fig. 1A) and the lateral superior olive, regions that do not contain CRF messenger RNA (ref. 14, and data not shown). Less pervasive perikaryal staining is also detected in the external plexiform layer of the olfactory bulb and scattered cells of the lateral hypothalamus. Labelled axons and terminals are distributed throughout many brain regions and the length of the spinal cord. There are prominent urotensin-like immunoreactive stained projection fields in the lateral septal nucleus (Fig. 1B) among other areas in which type-2 CRF receptors are expressed^{10,12}.

The CRF receptors characterized so far are encoded by two distinct genes and differ in their anatomical distribution and affinities for urotensin and sauvagine: type 1, or CRF-R1, was isolated from pituitary^{6,9} and brain^{7,8}, and type 2, or CRF-R2,

† To whom correspondence should be addressed.

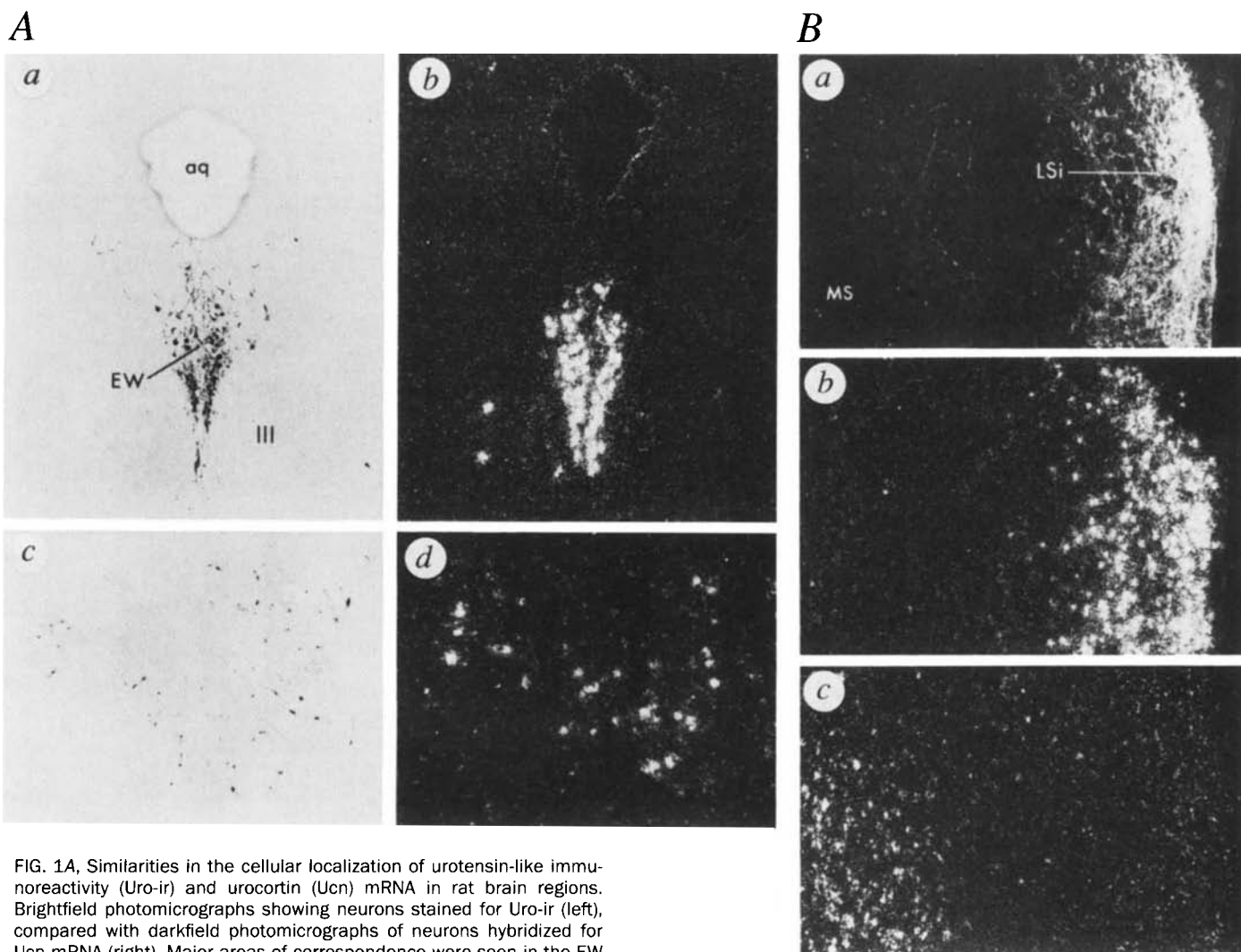


FIG. 1A, Similarities in the cellular localization of urotensin-like immunoreactivity (Uro-ir) and urocortin (Ucn) mRNA in rat brain regions. Brightfield photomicrographs showing neurons stained for Uro-ir (left), compared with darkfield photomicrographs of neurons hybridized for Ucn mRNA (right). Major areas of correspondence were seen in the EW (a, b) and the lateral superior olive (c, d). Although both sites have been found before to crossreact with CRF antisera^{14,29}, CRF mRNA is expressed in neither (data not shown). Moreover, immunostaining in both cell groups was eliminated by competition with 3 μ M Uro, but not by 300 μ M CRF. Magnifications: top, 135 $\times$; bottom, 150 $\times$. **B**, Congruence of a Uro-ir terminal field and cellular CRF-R2, but not CRF-R1, mRNA expression in the septal region of the rat brain. Darkfield photomicrographs showing that Uro-ir axons and terminals (a) and neurons expressing CRF-R2 mRNA (b) are similarly distributed in the intermediate part of the lateral septal nucleus (LSi), whereas cells expressing CRF-R1 mRNA (c) are preferentially sequestered in the medial septal nucleus (MS), a region lacking in substantial Uro-ir projections. aq, aqueduct; III, oculomotor nucleus. Magnifications, 50 $\times$. **METHODS**. All histochemistry was done on perfusion-fixed (4% paraformaldehyde) 30- μ m-thick frozen sections from adult male rats. Animals used for cellular localization of Uro-ir received intracerebroventricular (i.c.v.) injections of colchicine (50 μ g in 25 μ l) 72 h before being killed.

was cloned from brain¹⁰ and heart¹¹⁻¹³. Two splice variants of CRF-R2 (refs 10-13, 15) differ with respect to their amino-terminal domains and anatomical distribution. A 431-amino-acid CRF-R2 subtype (CRF-R2 β) is found in peripheral tissues, including the duodenum, skeletal muscle, epididymis and perivascular cells of the heart¹², as well as in brain, where it is associated with blood vessels¹⁵. A 411-amino-acid variant of this receptor (CRF-R2 α)¹⁰ is expressed in limited areas of the brain, including the lateral septal, ventromedial hypothalamic and medial amygdaloid nuclei, and displays a much more restricted distribution than that of CRF-R1¹⁶. Urotensin and sauvagine are more potent than CRF at increasing cAMP in cells transfected with either variant of CRF-R2 (refs 10-13), consistent

Immunostaining: An avidin-biotin-immunoperoxidase method³⁰ was used to localize an antiserum raised in rabbit against white-suckerfish Uro using methods previously described for CRF²⁷ (1:2,000 dilution; code, PBL 5712). Sections were stained in the presence of 30 μ M CRF. **Localization of mRNA:** *In situ* hybridization was performed as described³⁰ using ³⁵S-labelled antisense and sense (control) cRNA probes (1-3 $\times 10^9$ d.p.m. μ g⁻¹) under high-stringency conditions (50% formamide, with final washes in 0.2 $\times$ SSC at 70 $^{\circ}$ C). cRNA copies were synthesized from: (1) a 550-bp Ucn cDNA clone encompassing the entire coding region and $\sim$ 60 bp of 5' and 3' untranslated sequence; (2) a full-length (1.3-kb) rat CRF-R1 cDNA; or (3) a 1.0-kb mouse CRF-R2 β cDNA encompassing a portion of 5' untranslated region and 72% of coding sequence. Because CRF-R2 mRNA in brain¹⁰ is a smaller variant of the species found in heart¹¹⁻¹³, probes were hydrolysed to $\sim$ 250-nt fragments.

with the idea that a urotensin- or sauvagine-like molecule may be the preferred mammalian ligand for this receptor.

The latter view is reinforced by observations that in the septal region, the distribution of urotensin-like immunoreactivity is in register with patterns of CRF-R2, but not CRF-R1, expression (Fig. 1B). Terminals with this immunoreactivity end in pericellular basket-like arrays in the regions of the lateral septal nucleus that comprise the major central site of CRF-R2 expression. Cells expressing CRF-R1 mRNA are preferentially sequestered in the medial septal nucleus¹⁶, a region deficient in urotensin-like immunoreactivity.

A cDNA library constructed with mRNA derived from a rat midbrain dissection encompassing the EW was screened with an

TABLE 1 Binding and biological properties of urocortin and related peptides

a Functions assayed <i>in vitro</i>						
Peptide	ACTH secretion by anterior pituitary cells (EC ₅₀ , nM)	cAMP in cells expressing CRF-R1 (EC ₅₀ , nM)	cAMP in cells expressing CRF-R2β (EC ₅₀ , nM)	Binding to stable CHO CRF-R1 (K _i , nM)	Binding to stable CHO CRF-R2β (K _i , nM)	Binding to CRF-BP CRF-R1 (K _i , nM)
Urocortin (rat)	0.006 ± 0.003	0.8 ± 0.1	0.18 ± 0.04	0.16 (0.08–0.32)	0.41 (0.26–0.66)	0.10 (0.09–0.11)
CRF (rat)	0.043 ± 0.012	1.9 ± 0.8	1.7 ± 0.4	0.95 (0.47–2.0)	17 (10–29)	0.21 (0.17–0.26)
Urotensin I (suckerfish)	0.017 ± 0.003	3.1 ± 1.7	0.74 ± 0.1	0.43 (0.23–0.81)	3.0 (1.8–4.8)	0.028 (0.021–0.036)
Sauvagine (frog)	0.033 ± 0.010	2.5 ± 1.4	0.5 ± 0.2	1.2 (0.54–2.5)	2.0 (1.1–3.6)	65 (46–92)
b Effects on mean arterial pressure						
Urocortin dose (μ kg ⁻¹)	Δ MAP (mm Hg)	Duration (min)				
Control	-1.2 ± 0.1	2.3 ± 0.5				
0.94	-5.5 ± 0.3	10.9 ± 0.8				
1.89	-10.1 ± 0.5	34.1 ± 2.1				
3.77	-18.3 ± 0.7	94.5 ± 3.2				
18.85	-26.9 ± 1.8	215.4 ± 5.3				
Peptide	Δ MAP (mm Hg)	Duration (min)				
Urocortin	-18.3 ± 0.7†	94.5 ± 3.2				
CRF	-5.8 ± 0.5*	50.2 ± 4.6				
Urotensin I	-8.8 ± 0.7*†	78.6 ± 3.5				
c Effects on plasma ACTH concentration <i>in vivo</i>						
Peptide	ACTH (pg ml ⁻¹)				Area under curve (pg ml ⁻¹ h)	
	10 min	30 min	60 min	120 min		
Control	37 ± 23	38 ± 16	22 ± 12	25 ± 12	55 ± 9	
CRF (1 μg kg ⁻¹)	325 ± 69	238 ± 15	126 ± 33	49 ± 5	300 ± 29	
Urocortin (1 μg kg ⁻¹)	408 ± 70	659 ± 53	405 ± 51	103 ± 11	732 ± 40‡§	
CRF (5 μg kg ⁻¹)	339 ± 61	422 ± 66	275 ± 75	82 ± 23	510 ± 106‡	
Urocortin (5 μg kg ⁻¹)	430 ± 36	994 ± 81	935 ± 80	260 ± 38	1354 ± 66‡§	

a, Comparison of the binding properties and functional activities of synthetic urocortin with CRF, urotensin and sauvagine *in vitro*. ACTH *in vitro* bioassay: Rat anterior pituitary cells were established in culture and treated with test peptides as described²⁷. Secreted ACTH was measured using a kit (Diagnostic Products) and *in vitro* potencies were determined using the Allfit computer program with results expressed as the average ± s.e.m. of three replicate bioassays. cAMP accumulation assay: COS-M6 cells were transfected with rat CRF-R1⁷ or mouse CRF-R2β¹², treated with test substances, extracted and cAMP determined⁶. EC₅₀ (± s.e.m.) values were determined from 3–8 independent experiments for each test peptide. CRF-receptor-binding assay: The affinities of test peptides for CRF-R1⁶ and CRF-R2β¹² stably expressed in CHO cells were determined by competitive displacement of [¹²⁵I](Nle21, Tyr32) ovine CRF (for CRF-R1) or of [¹²⁵I]-Tyr⁹Ucn (for CRF-R2β) as described⁶. Data from at least 3 experiments were pooled and inhibitory dissociation constant (K_i) values (95% confidence limits) were calculated using the LIGAND program. CRF-BP ligand-binding assay: The affinity of recombinant human CRF-BP for test peptides was measured on the basis of the displacement of [¹²⁵I]-D-Tyr⁹hCRF as previously described²⁸. b, Effect of graded doses of urocortin on mean arterial pressure (MAP) and duration of hypotension in the rat (top) and relative abilities of urocortin, CRF and urotensin to reduce MAP (bottom). Data are presented as mean ± s.e.m., n = 5. Data were analysed by one-way ANOVA followed by Dunn's test (StatView 4.0). Responses to urocortin differed significantly from those to control infusions at all doses tested (P < 0.005). Mean arterial pressure was calculated as: 2/3 diastolic pressure + 1/3 systolic pressure. A reflex tachycardia (+20 beats per min) in response to 3.77 μg kg⁻¹ urocortin was observed; no significant changes in pulse pressure were recorded at any dose tested. Measurements of arterial pressure: Experiments were carried out on individually housed, awake, freely moving Sprague-Dawley albino rats 24 h after implantation of femoral arterial and jugular venous catheters for monitoring of arterial pressure and infusion of peptide, respectively. 1–1.5 h before the experiment, the arterial catheter was connected to a blood-pressure transducer (Statham) and the venous catheter to an infusion pump (Bioanalytical Systems). Heart rate was computed from the blood pressure pulses using a biotachometer (Gould). The peptides (0.02 μg μl⁻¹) in saline were infused intravenously in randomized sequence at a rate of 50 μl min⁻¹ over 15 s–5 min, depending on the dose (1 μg min⁻¹). Infusion of vehicle and heat-denatured peptide (3.77 μg kg⁻¹ urocortin boiled for 30 min) were used as control procedures. c, Effect of graded doses of urocortin on plasma ACTH concentration in the rat. Data are expressed as means ± s.e.m., n = (3–4). The area under the curve was calculated by integration of plasma ACTH concentrations over the 2-h-test period and analysed by one-way ANOVA followed by Dunn's test (Statview 4.0). ACTH bioassay: Experiments were done under the same conditions as in b, 48 h after implantation of jugular venous catheters for withdrawal of blood samples and injection of peptides. Plasma ACTH concentrations were determined using a two-site immunoradiometric assay (Nichols Institute, San Juan, Capistrano, CA). 3 h after connection of catheters to sampling syringes, an initial 0.5-ml blood sample was taken (ACTH values similar for all groups: means, 9–24 pg ml⁻¹). Either vehicle (PBS containing 0.1% BSA, control) or peptide (0.2 ml per 100 g) was injected intravenously over 0.5 min. Subsequent blood samples were taken and 0.3 ml heparinized saline administered as replacement fluid.

* Differs significantly from urocortin response (P < 0.005).

† Differs significantly from CRF response (P < 0.01).

‡ Significantly different from control (P < 0.001).

§ Significantly different from CRF response at the same dose of peptide (P < 0.0001).

urotensin probe. The largest of four independent clones is a full-length cDNA containing a single open reading frame that encodes a protein deduced to be 122 amino acids long with an N-terminal methionine and consensus signal-peptide sequence (Fig. 2a). The carboxy terminus includes a putative 40-amino-acid peptide flanked by proteolytic processing sites. The C-terminally amidated putative peptide clearly belongs to the CRF

family, which includes fish urotensin, frog sauvagine and insect diuretic hormones³. We call this peptide urocortin (Ucn) to reflect its relationship (Fig. 2b) to urotensin (63% identity with suckerfish urotensin) and to CRF (45% identity with rat CRF), and because it possesses biological activities of both peptides. Urocortin also shares primary structural features with sauvagine (35% identity), especially in the N-terminal portion (Fig. 2b).

FIG. 2 a, Deduced amino-acid sequence encoding urocortin. Amino acids are numbered starting with the initiating methionine. The Ucn peptide sequence (boxed) is immediately preceded by the proteolytic processing site (R-R) and followed by amino acids (G-K) presumed to be involved in C-terminal amidation and cleavage from the precursor. The full nucleotide sequence has been deposited in the GenBank data base (accession no. U33935). b, Comparison of the amino-acid sequence of rat urocortin (rUcn) with white-suckerfish urotensin I (sfUro), rat corticotropin-releasing factor (rCRF), frog sauvagine (SvG) and *Manduca sexta* diuretic hormone (Mas-DH³). Identical amino acids are shaded. Sequences are listed from highest to lowest homology scores, determined using the Lasergene System alignment program with the Clustal algorithm and an identity residue weight table. Square denotes C-terminal amidation.

METHODS: Urotensin

probe: Two oligonucleotides corresponding to amino acids NDPPIS (sense: 5'-AAGCAGCATCTCCGATCTCC-3') and LDEVGK (antisense: 5'-CTTCCAACCTCATCTAG-3') of goldfish (*Carassius auratus*) Uro (M. Stenzel-Poore and W. V., unpublished results), were used to prime 200 ng genomic DNA by PCR amplification (35 cycles of 94 °C for 1', 60 °C for 1', 72 °C for 1'). The insert was subcloned into pBluescript, the sequence verified and the plasmid DNA was used as template with the sense and antisense primers to generate a 129-bp [α -³²P-dCTP]-labelled Uro probe by PCR amplification (40 cycles of 94 °C for 1', 55 °C for 2', 72 °C for 3'). Library construction: Total RNA was prepared by guanidine thiocyanate extraction of a rat midbrain block encompassing the EW from 50 male rats and poly (A⁺) mRNA was obtained with oligotex-dT (Qiagen). The corresponding cDNA

b

-DDPPLSIDLTFHLLRT	LLELARTQSQ	-RERAEQNRIIFDSV	■	rUcn
NDPPPTSIDLTFHLLRN	MIEMARNENQ	-REQAGLNRYLDEV	■	sfUro
SEEPPISLDLTFHLLRE	VLEMARAEQL	-AQQAHSNRKLMET	■	rCRF
-EGPPIISIDLSLELLRKM	IEIEKQEKE	-KQQAANNRLLLDIT	■	SvG
-RMPISLSIDLPM SVLRQ	KLSELEKERVHALLRA	AANRNFLNDI	■	Mas-DH

was oligo(dT)-primed and ligated into λ gt10 (Stratagene) using *EcoRI*/*NotI* adaptors yielding a library of 6×10^6 recombinants. Library screening: About 0.6×10^6 phage plaques were screened by hybridization, using $\sim 1 \times 10^6$ c.p.m. ml⁻¹ of the 129-bp Uro probe at 42 °C in standard buffers containing 20% formamide. Final washes at 42 °C were in 2 $\times$ SSC/0.1%SDS. Phage DNA was purified from the four independent clones obtained and subcloned into pBluescript (Stratagene). Both strands were sequenced using the Sequenase kit (US Biochemical).

The precursor portion of urocortin is half the length of urotensin¹⁷ and CRF precursors¹⁸, with little sequence similarity to either.

Northern blot analysis of poly (A⁺) mRNA from microdissected midbrain blocks encompassing the EW nucleus shows a single transcript of ~ 1 kilobase (kb), which agrees with the size of the cloned cDNA sequence (not shown). The urotensin antiserum used for immunological screening binds the radiolabelled N-terminal analogue [¹²⁵I-Tyr⁹]Ucn with high affinity and capacity, and urotensin/urocortin immunoreactivity can be detected in extracts of rat central nervous system (CNS) and duodenum by a heterologous radioimmunoassay (urotensin antiserum/urocortin tracer/urocortin standard). The predominant immunoreactive form in partially purified rat CNS tissue elutes at a similar position to synthetic urocortin in gel filtration chromatography, consistent with the idea that studies with anti-urotensin sera are measuring urocortin cleaved from its precursor.

Urocortin induces accumulation of intracellular cyclic AMP in COSM6 cells transiently transfected with either CRF-R1 or CRF-R2 β (Table 1a). The concentration for half-maximal efficacy (EC₅₀) of urocortin on CRF-R1 is slightly lower than that for CRF. By contrast, urocortin is almost tenfold more potent than CRF at increasing cAMP in cells expressing CRF-R2 β . Urocortin binds six times more strongly than CRF to membranes from cells expressing CRF-R1, and 40 times more strongly to expressed CRF-R2 β . A radiolabelled urocortin analogue, [¹²⁵I-Tyr⁹]Ucn, is superior for monitoring CRF-R2 than are tracers derived from CRF analogues.

The broad CNS distribution of a 37 kD CRF binding protein (CRF-BP) which binds to and inactivates CRF led us to propose that CRF-BP may function to limit the action of endogenous CRF within the brain¹⁹. Because of the high affinity of the binding protein for Ucn (Table 1a), it is possible that CRF-BP may similarly modulate the activity of this new mammalian member of the CRF family.

Synthetic urocortin is a potent stimulator of ACTH secretion from rat pituitary cells in primary culture and *in vivo* in the unanaesthetized rat (Table 1), which would be expected from

the peptide's ability to activate CRF-R1, the receptor subtype expressed by corticotropes¹⁶. Although it is not yet known whether urocortin gains access to the pituitary through the pituitary portal system or through the peripheral blood, the residual secretion of ACTH in female mice with a disrupted CRF gene²⁰ suggests that other secretagogue(s) like urocortin might maintain ACTH secretion in the absence of CRF.

Urocortin mRNA accumulates in the EW nucleus and the lateral superior olive, sites that are coincident with dominant cellular sites of urotensin-like immunoreactivity (Fig. 1). Weaker signals are detected over large, presumably somatomotor, neurons of the facial, hypoglossal and ambigular motor nuclei, over a subset of magnocellular neurosecretory neurons of the supraoptic nucleus, and over cells distributed diffusely in the caudal lateral hypothalamic area. The EW nucleus is best known as the origin of parasympathetic preganglionic innervation of the ciliary ganglion, although in rats it contributes in only a minor way to the pathway innervating the anterior eye segment²¹. There are descending projections to the brain stem and spinal cord arising from the EW nucleus²². The lateral superior olive functions principally in the processing of auditory information, and its major connections are with similarly dedicated structures in the midbrain and pons²³. The source of the urotensin-like immunoreactivity in the forebrain, for example in the lateral septal nucleus, is unknown, although projections to the septal region from the lateral hypothalamus have been described²⁴. Further histochemistry is needed to clarify the distribution and origins of urotensin/urocortin-like immunoreactive projection fields in the CNS.

We tested the ability of urocortin to activate cells bearing type-2 CRF receptors *in vivo* assaying for expression of the transcription factor Fos, induced by central administration of the peptide. Fos is a broadly applicable generic marker of synaptic and/or transcriptional activation²⁵. The pattern of urocortin-induced Fos expression is distinct from that following similar injections with CRF (J.B. and P.E.S., unpublished observations), and includes activation of CRF-R2 expression at several major sites including the lateral septal, dorsal raphe (Fig. 3) and

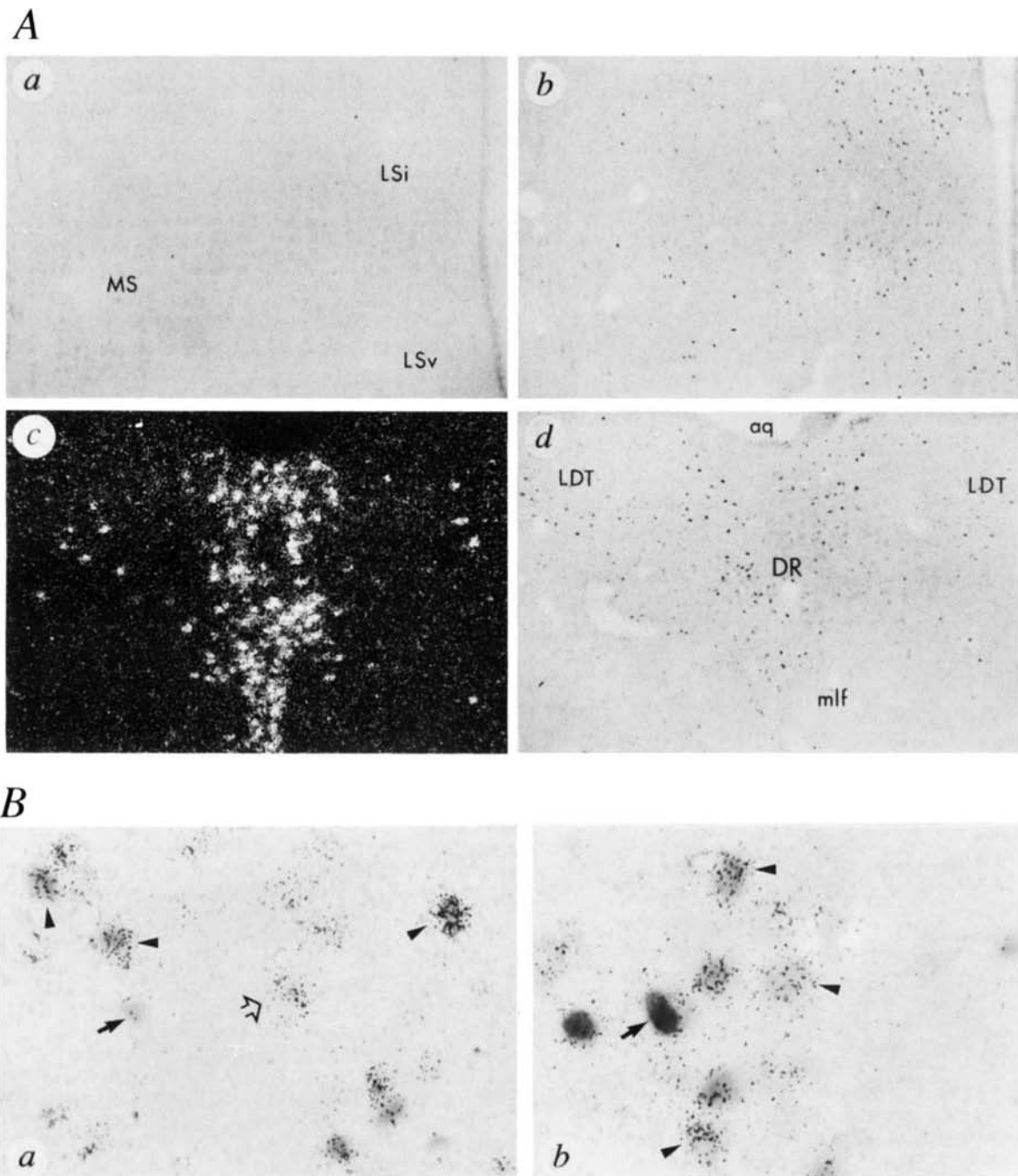


FIG. 3 A, Centrally administered Ucn induces Fos at the sites of CRF-R2 mRNA expression. Patterns of nuclear Fos induction seen in the septal region 2 h after i.c.v. injection of vehicle (a) or 1 μ g Ucn (b); Ucn-induced Fos activity is preferentially distributed in the intermediate part of the lateral septal nucleus (LSi), a major locus of CRF-R2 expression (compare Fig. 1b) and, to a much lesser extent, in the type-1 receptor-enriched medial septal nucleus (MS). The dorsal raphe nucleus (DR), another major site of CRF-R2 (c), but not CRF-R1²⁰, mRNA expression, also gave a focally dense Ucn-induced Fos response (d). Other abbreviations: LDT, laterodorsal tegmental nucleus; LSv, ventral lateral septal nucleus; mlf, medial longitudinal fasciculus. Magnification $\times 75$. B, Localization of Ucn-induced Fos to neurons expressing CRF-R2 mRNA. Brightfield photomicrographs of sections through the lateral septal (a) and dorsal raphe (b) nuclei showing cells with Ucn-induced nuclear Fos activity (grey) and decorated with silver grain clusters (black

dots), indicating a positive CRF-R2 mRNA signal (arrowheads). Cells labelled singly for Fos (black arrows) or the type-2 receptor transcript (open arrows) are shown for contrast. Magnification, $\times 750$.

METHODS. Adult male rats implanted stereotactically with guide canulae for i.c.v. injection at least 7 days before experimentation were given 1 μ g synthetic Ucn or saline vehicle in 10 μ l and killed 2 h later, the time at which maximal Fos induction was observed. Frozen sections of the brain were collected as for Fig. 1 and prepared for concurrent dual localization of Fos activity and CRF-R2 mRNA, as described³⁰. Fos activity was detected using a polyclonal antiserum raised against a synthetic N-terminal fragment of human Fos (Oncogene Sciences), diluted 1:2,000 for application to tissue sections. Staining was abolished by preadsorbing the antiserum overnight at 4 $^{\circ}$ C with 50 μ M synthetic immunogen.

solitary tract nuclei. Urocortin-stimulated cellular activation is not restricted to CRF-R2-expressing neuronal populations, as a strong Fos response is also seen in regions enriched in CRF-R1 receptors (for example, pontine gray, nucleus incertus), as well as in cell groups that do not express either CRF-receptor subtype (such as the central nucleus of the amygdala, locus coeruleus).

In keeping with the distribution of urocortin and its receptors in the brain, the neuropeptide has been found (G. Koob and M. Spina, personal communication) to have behavioural actions following intracerebroventricular administration; one notable example is that urocortin profoundly suppresses feeding behaviour in fasted rats.

Urocortin, when given intravenously in freely moving rats, yields a pronounced and prolonged reduction of mean arterial blood pressure (MAP), an effect reported earlier for urotensin and CRF^{2,26}. Both the size and duration of this effect are dose-dependent and greater than those elicited by CRF or urotensin (Table 1b). The perivascular localization of CRF-R2 β in the heart¹⁰, and the greater potency of sauvagine and urotensin in lowering blood pressure and dilating the superior mesenteric arterial bed²⁶, suggest that urocortin is highly effective in this CRF-R2 β -mediated action. Our observation of urocortin-like immunoreactivity in duodenal extracts as well as the broad distribution of one of its proposed cognate receptors may be indicative of an extensive peripheral urocortin system in addition to its discrete organization in the brain. □

Received 10 July; accepted 22 September 1995.

- Vale, W., Spiess, J., Rivier, C. & Rivier, J. *Science* **213**, 1394–1397 (1981).
- Vale, W. et al. *Chemical and Biological Characterization of Corticotropin Releasing Factor in Recent Progress in Hormone Research*, Vol. 39, 245–270 (Academic, San Diego, 1983).
- Lederis, K., Fryer, J. N., Okawara, Y., Schonrock, C. & Richter, D. in *Fish Physiology* (ed. Farrell, A.) 67–100 (Academic, San Diego, 1994).
- Okawara, Y. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8439–8443 (1988).
- Stenzel-Poore, M. P., Heldwein, K. A., Stenzel, P., Lee, S. & Vale, W. W. *Molec. Endocr.* **6**, 1716–1724 (1992).
- Chen, R., Lewis, K. A., Perrin, M. H. & Vale, W. W. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8967–8971 (1993).
- Perrin, M. H., Donaldson, C. J., Chen, R., Lewis, K. A. & Vale, W. W. *Endocrinology* **133**, 3058–3061 (1993).

- Chang, C., Pearce, R., O'Connell, S. & Rosenfeld, M. G. *Neuron* **11**, 1187–1195 (1993).
- Vita, N. et al. *FEBS Lett.* **335**, 1–5 (1993).
- Lovenberg, T. W. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 836–840 (1995).
- Kishimoto, T., Pearce, R. V. II, Lin, C. R. & Rosenfeld, M. G. *Proc. natn. Acad. Sci. U.S.A.* **92**, 1108–1112 (1995).
- Perrin, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 2969–2973 (1995).
- Stenzel, P. et al. *Molec. Endocr.* **9**, 637–645 (1995).
- Imaki, J., Imaki, T., Vale, W. & Sawchenko, P. E. *Brain Res.* **547**, 28–36 (1991).
- Lovenberg, T. W., Chalmers, D. T., Liu, C. & De Souza, E. B. *Endocrinology* **136**, 3351–3355 (1995).
- Potter, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 8777–8781 (1994).
- Ishida, I., Ichikawa, T., & Deguchi, T. *Proc. natn. Acad. Sci. U.S.A.* **83**, 803–812 (1986).
- Thompson, R. C., Seasholtz, A. F. & Herbert, E. *Molec. Endocr.* **1**, 363–370 (1987).
- Potter, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4192–4196 (1992).
- Muglia, L., Jacobson, L., Dikkes, P. & Majzoub, J. A. *Nature* **373**, 427–432 (1995).
- Loewy, A. D. in *Central Regulation of Autonomic Functions* (eds Loewy, A. D. & Spyer, K. M.) 268–285 (Oxford Univ. Press, New York, 1990).
- Loewy, A. D., Saper, C. B. & Yamodis, N. D. *Brain Res.* **141**, 153–159 (1978).
- Webster, W. R. in *The Rat Nervous System* (ed. Paxinos, G.) 797–831 (Academic, San Diego, 1995).
- Saper, C. B., Swanson, L. W. & Cowan, W. M. J. *comp. Neuro.* **183**, 689–706 (1979).
- Morgan, J. I. & Curran, T. A. *Rev. Neurosci.* **14**, 421–451 (1991).
- Lederis, K. et al. *Rec. Prog. Horm. Res.* **41**, 553–576 (1985).
- Vale, W. et al. *Meth. Enzym.* **103**, 565–577 (1983).
- Sutton, S. W. et al. *Endocrinology* **136**, 1097–1102 (1995).
- Chung, R. Y., Mason, P., Strassman, A., Maciewicz, R. *Neurosci. Lett.* **83**, 13–19 (1987).
- Chan, R. K. W., Brown, E. R., Ericsson, A., Kovács, K. J. & Sawchenko, P. E. *J. Neurosci.* **13**, 5125–5138 (1993).

ACKNOWLEDGEMENTS. We thank K. Lederis for urotensin and sauvagine antisera used initially to identify related immunoreactive substances in rat tissue extracts; S. Lahrichi and R. Kaiser for synthetic peptides; J. Black for radioimmunoassays; A. Corrigan for *in vitro* ACTH-release bioassays; and L. Bilezikjian for critically reading the manuscript. This work was supported in part by DK-26741, the Foundation for Research, the Foundation for Medical Research, the Kleberg Foundation, the Adler Foundation, the Hearst Foundation, and the Brazilian CNPq.

A stomatin-like protein necessary for mechanosensation in *C. elegans*

Mingxia Huang*, Guoqiang Gu, Edwin L. Ferguson* & Martin Chalfie†

Department of Biological Sciences, 1012 Fairchild, Columbia University, New York, New York 10027, USA

The *mec-2* gene is required for the function of a set of six touch receptor neurons in the nematode *Caenorhabditis elegans*; *mec-2* mutants, which are touch-insensitive, have touch cells that appear morphologically normal^{1,2}. Gene interaction studies suggest that *mec-2* positively regulates the activity of the putative mechanosensory transduction channel (ref. 3 and the present paper), comprised in part of proteins encoded by the two degenerin genes *mec-4* and *mec-10* (refs 3–5). The central region of the *mec-2* protein (MEC-2) is very similar to stomatin, an integral membrane protein (band 7.2b) in human red blood cells that is thought to regulate cation conductance⁶. MEC-2–LacZ fusions are distributed along the touch receptor axons. This axonal distribution, which is mediated by the *mec-2*-specific amino terminus, is disrupted by mutations in *mec-12*, an α -tubulin gene needed for touch cell function. Our results indicate that MEC-2 links the mechanosensory channel and the microtubule cytoskeleton of the touch receptor neurons. Such linkage provides the basis for a mechanism of mechanosensation whereby microtubule displacement leads to channel opening.

mec-2 was cloned by transposon tagging and transformation rescue of the touch-insensitive (*Mec*) phenotype (Fig. 1a). The full-length *mec-2* complementary DNA encodes an open reading frame of 481 amino acids (Fig. 1b). The central 247 amino acids of MEC-2 share extensive sequence similarity with stomatin (65% identity and 85% similarity; Fig. 1b). Both proteins are very hydrophilic, except for a single hydrophobic domain. The similarity between MEC-2 and stomatin starts with the hydro-

phobic domain and extends to the carboxyl end of the human protein.

Stomatin is lost in red blood cells of patients with hereditary stomatocytosis, an autosomal dominant haemolytic anaemia (reviewed in ref. 7). The swelling and lysis of red blood cells in stomatocytosis has led to the suggestion that stomatin regulates membrane conductance^{6,7}. The striking sequence similarity between stomatin and MEC-2 suggests that MEC-2 has a similar function and subcellular location. This hypothesis is consistent with our finding that *mec-2* is needed for function of the putative mechanosensory channel in the touch cells³. The localization of stomatin^{6,7} in the membrane suggests the possibility of interactions of stomatin and MEC-2 with ion channels. The roles of these proteins may not be identical. Stomatin is thought to inhibit membrane conductance, whereas *mec-2* seems to regulate the putative degenerin channel positively³. These opposite effects may arise from differences between stomatin and MEC-2 (for example, the N and C termini specific to MEC-2), or between the molecules that interact with them. Alternatively, other defects may be present in patients with stomatocytosis.

We identified both missense and nonsense mutations in the stomatin-like region of *mec-2*, confirming the identity of the gene. These *mec-2* mutations vary in expressivity and some exhibit complex intragenic complementation patterns (ref. 1; M.H. and M.C., unpublished data), suggesting that MEC-2 forms multimers or interacts with other proteins. As stomatin binds to the red blood cell cytoskeleton through its carboxyl region (G. Stewart and J. Sinard, personal communication), the similar region in MEC-2, which is highly conserved, may have an analogous function.

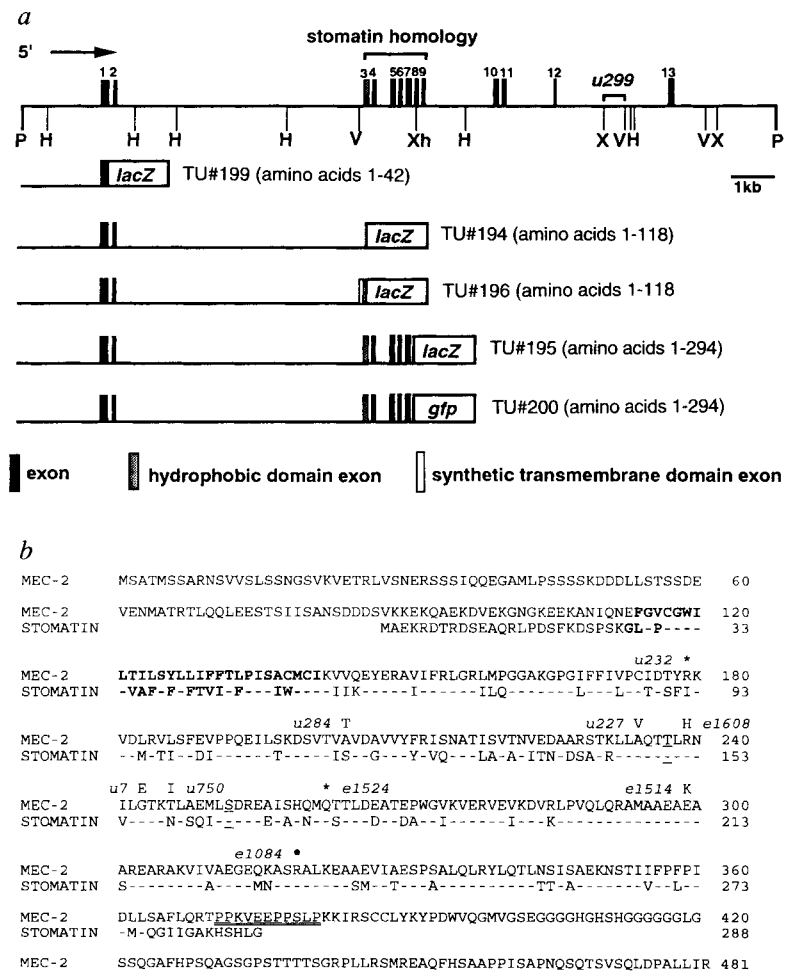
MEC-2 extends beyond the stomatin-like sequence at both its N and C termini. As described below, the MEC-2-specific N terminus may interact with microtubules. The MEC-2-specific C terminus may also bind to other proteins, as it contains a proline-rich region (amino acids 371 to 381) that has characteristics of a SH3-binding site, a site involved in binding to either signal-transduction or cytoskeletal proteins (reviewed in ref. 8). The degenerin-like, mammalian epithelial sodium channel (ENaC) proteins also contain SH3-binding sites that bind to cytoskeletal proteins⁹, but the *C. elegans* degenerins do not. We do not know whether the proline-rich region in MEC-2, a protein needed for degenerin function, serves a similar function to the ENaC domain.

* Present addresses: Department of Biochemistry, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA (M.H.); Department of Molecular Genetics and Cell Biology, University of Chicago, 920 E. 58th Street, Chicago, Illinois 60637, USA (E.L.F.).

† To whom correspondence should be addressed.

FIG. 1 Genomic structure and protein sequence of *mec-2*. **a**, Restriction map, intron/exon structure, and reporter fusion constructs of *mec-2*. Horizontal arrow indicates the direction of transcription. Numbered boxes along the top line show the exon positions. Exons 3 to 9 encode sequences that share similarity with stomatin. Exon 3 encodes the hydrophobic domain. Also bracketed is the approximate position of the 1.5-kb insertion in *mec-2(u299)*. All fusions were made using *mec-2* genomic DNA and fusion sites for *mec-2-lacZ* and *mec-2-gfp* constructs are indicated. MEC-2 amino acids encoded by the fusions are given in parentheses. TU#196 is identical to TU#194 except that it contains a single synthetic transmembrane domain¹⁰ between MEC-2 and β -GAL. Abbreviations: H, *HindIII*; P, *PstI*; V, *EcoRV*; X, *XbaI*; Xh, *XhoI*. **b**, Alignment between the predicted protein sequences of MEC-2 and stomatin. Dashes in stomatin represent identity to MEC-2 sequence. The hydrophobic domains are indicated in bold. Two conserved potential protein kinase C (PKC) phosphorylation sites are underlined. The potential SH3-binding site is doubly underlined. The position of several *mec-2* mutations (allele names are in italics) are indicated above the MEC-2 sequence. An asterisk indicates a nonsense mutation. The GenBank accession numbers for the *mec-2* genomic and cDNA sequences are U26736 and U26735, respectively.

METHODS. General methods are described in ref. 29. *u299* is a transposon insertion allele of *mec-2*. A Tc1-containing fragment that cosegregated with the *u299* mutation was isolated from *u299* strain and used to position the *mec-2* gene on the *C. elegans* physical map³⁰. Because this insertion was not lost in two revertants of *u299*, *u412* and *u413*, it is not the cause of the *u299* mutation. Two λ clones encompassing ~25 kb of genomic DNA around the site of the Tc1 insertion, however, identified a second polymorphism in *u299* animals that was missing in the revertant animals. Furthermore, three overlapping cosmids at *mec-2* locus (C49H2, W05G12 and C12H10) revealed a 1.5-kb upshift of a 6.5-kb *EcoRV* fragment on Southern blots in *u299* but not in wild type or revertants. The insertion is also seen with the 2.5-kb *XbaI* fragment. W05G12 also rescued the *mec-2(e1084)* mutant phenotype in germline transformation using a *rol-6* marker³¹. The 6.5-kb *EcoRV* fragment was used to screen two mixed-stage *C. elegans* cDNA libraries (C. Martin and M.C., unpublished data; ref. 32). All cDNA clones corresponded to the same transcript. The longest ones were 1.8 kb and appeared to contain the complete coding sequence of *mec-2*, with the longest open reading frame preceded by an in-frame stop codon. The *mec-2*-rescuing activity was located to a 20.5-kb *PstI* fragment, which contained 2.6 kb of DNA upstream of the start of the open reading frame. This 20.5-kb *PstI* fragment was subcloned into pBluescript



We examined *mec-2* expression using both *lacZ*¹⁰ and the gene encoding the *Aquorea victoria* green fluorescence protein (*gfp*)¹¹ as reporters (Figs 1, 2 and Table 1). β -Galactosidase activity was detected in the six touch receptor neurons and a few other cells, particularly in the head (Fig. 2a). As a slight chemotaxis defect has been found in animals with the *mec-2(e75)* mutation^{12,13}, the cells in the head may be important for chemotaxis. GFP fluorescence was seen in the same cells, although fluorescence was higher in the six touch cells than in the other cells (data not shown). *mec-2* expression in the touch cells depends on *unc-86* and *mec-3*, two homeobox genes that control touch-cell differentiation^{14,15} (Fig. 2c, d and Table 1).

The extensive sequence similarity of MEC-2 and stomatin suggests that MEC-2 is also an integral membrane protein. Expression of β -galactosidase from *mec-2-lacZ* fusions suggests that both the N and C termini of MEC-2 are cytoplasmic. Fusions made before (TU#194) or after (TU#195) the hydrophobic region produce active β -galactosidase. However, when a synthetic transmembrane domain is added to the hydrophilic N terminus of MEC-2 in (TU#196), there is no activity. As extracellular β -galactosidase is inactive in *Escherichia coli*¹⁶ and *C. elegans*¹⁰, these data suggest a monotopic topology for MEC-2 in the membrane. A similar topology has been proposed

for stomatin^{7,17}. The predicted secondary structure¹⁸ for the single hydrophobic domain of both MEC-2 and stomatin favours β -sheet over α -helix. This domain may span the membrane twice, with a bent structure of two antiparallel β -strands. Fusions with at least the first 118 amino acids of MEC-2 resulted in β -galactosidase activity in both the cell bodies and the processes of the touch cells (Fig. 2a and Table 1). In contrast, a fusion with only the N-terminal 42 amino acids of MEC-2 (from TU#199) resulted only in staining of the cell bodies (Fig. 2b and Table 1). A few PLM cells (Table 1) showed process staining in these animals, but this staining was never as extensive as with the longer fusions. These data indicate that the N terminus of MEC-2 contains a region that is sufficient for the axonal distribution of the fusion products.

The most prominent components of the touch cell cytoskeleton are the large-diameter, 15-protofilament microtubules that are unique to these cells and fill the touch-cell axons¹⁹. These microtubules, which form the *mec-7* β -tubulin^{1,20} and the *mec-12* α -tubulin (ref. 2, and M. Hamelin, M. Chou and J. Culotti, personal communication), are needed for mechanosensation but not for process outgrowth¹⁹. Mutations that eliminate these microtubules (for example, *mec-7(e1506)* and *mec-12(e1607)*)^{2,21} result in touch-insensitive animals possessing touch cells with

TABLE 1 Expression of MEC-2-LacZ fusions

Fusion	Mutation	N	ALM staining			Cell body	PLM staining			Cell body
			Per cent of total	Complete	Partial		Per cent of total	Complete	Partial	
TU # 194A	—	80	93	57	27	15	99	69	13	18
TU # 194B	—	82	91	85	7	9	93	80	9	11
TU # 199	—	50	57	0	0	100	78	7	3	90
TU # 194A	<i>mec-3(u6)</i>	60	1	0	0	100	66	0	0	100
TU # 194A	<i>unc-86(e1416)</i>	45	0	—	—	—	0	—	—	—
TU # 194A	<i>mec-7(e1506)</i>	27	85	41	2	57	93	14	20	66
TU # 194B	<i>mec-7(e1506)</i>	53	88	42	9	49	91	7	22	70
TU # 194B	<i>mec-12(e1607)</i>	78	92	61	14	25	90	53	9	38
TU # 194A	<i>mec-12(u63)</i>	68	88	12	23	65	94	7	10	83
TU # 194A	<i>mec-12(u241)</i>	92	68	11	28	61	92	2	18	80
TU # 195A	—	46	82	40	29	30	82	20	50	29
TU # 195B	—	36	63	29	49	22	89	8	16	76
TU # 195A	<i>mec-12(u63)</i>	62	88	2	19	79	87	1	2	97
TU # 195A	<i>mec-12(u241)</i>	58	88	0	18	82	89	0	3	97

For the TU #194 and TU #195 fusions two separate transgenic lines (A and B) were examined. Not all cells stain because the extragenic arrays of the fusions are sometimes lost somatically. N, number of animals examined. Per cent of total, proportion of cells at a given time that had any staining. Complete, partial, and cell body numbers are percentages of the stained cells that had processes that extended to the pharynx (ALM) or past the PVM cell (PLM; see Fig. 2), past the AVM cell but not to the pharynx (ALM) or to the PVM cell (PLM), and not past the AVM cell (ALM) or anus (PLM), respectively.

normal-length processes; axon outgrowth occurs because the large microtubules are replaced by a smaller number of smaller-diameter microtubules¹⁹.

To test whether microtubules are needed for the distribution of the larger MEC-2-LacZ fusions, we examined this distribution in *mec-7* and *mec-12* mutants (Fig. 2 and Table 1). The *mec-7(e1506)* and *mec-12(e1607)* mutations had a modest effect on the distribution of the TU #194 fusion. These results suggest that the large microtubules are needed for optimal distribution of the fusion, but they can be partially replaced by the smaller microtubules (the partial effect may be due to the smaller number of the latter microtubules in the mutants¹⁹). In contrast, two non-null mutations in *mec-12*, *u63* and *u241* strongly affected the axonal distribution of both the TU #194 and TU #195 fusions. Mutant cell bodies stained with the same intensity as wild-type, but axons stained less intensely and essentially only near the cell body (Fig. 2e, f) (the neuronal processes are of normal length in these mutants; data not shown). Both *u63* and *u241* are *mec-12* missense mutations. The recessive *u63* allele, which is unusual in not preventing the formation of the 15-protofilament microtubules², has a Glu415Lys change in the C terminus of the α -tubulin. This defect is three residues from the region needed for MAP1 and MAP2 binding in other α -tubulins^{22,23}. The semi-dominant *u241* allele results in the disappearance of the 15-protofilament microtubules in the touch-cell processes² and has a Gly354Glu change near the C terminus. These data, particularly

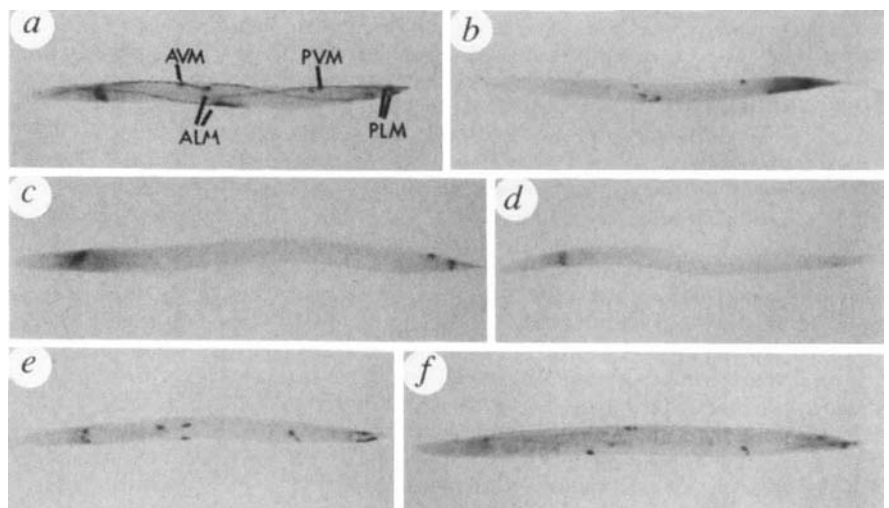
the results with *u63*, suggest that the N terminus of MEC-2 interacts with microtubules or their associated proteins.

Genetic interaction studies suggest that *mec-2*, *mec-7* and *mec-12* are needed for the proper function of the degenerin channel. Specifically, mutations in *mec-2* and *mec-12* suppress a degeneration-causing mutation in *mec-10* (fewer animals have degenerating touch cells)³, and otherwise recessive mutations in *mec-2*, *mec-7* and *mec-12* dominantly enhance the touch-insensitive phenotype of a temperature-sensitive allele of *mec-4* (G. Gu and M.C., details to be published elsewhere). In a screen for new dominant enhancing mutations of this temperature-sensitive mutation, *mec-4(u45)* (ref. 2), we identified a *mec-2* mutation, *u750*, that strongly enhances the *mec-4(u45)* phenotype (85% of the animals are touch-insensitive instead of <1%; $n > 200$). The *u750* allele contains a missense mutation in the stomatin-like region of *mec-2* (Fig. 1b).

These results lead us to suggest that MEC-2 links, directly or indirectly, the degenerin channels and the underlying microtubules. The arrangement of the 15-protofilament microtubules within the touch cells is consistent with this hypothesis. In adult cells the microtubules are short (20 μ m) compared with the length of the cell process (~400 μ m), and are distributed along its length¹⁹. As animals are touch-sensitive at all points along the process¹, the entire axon is likely to act as a sensory receptor. In cross-sections, 40–60 microtubules form a bundle that fills the process²⁴. The distal ends of the microtubules are always on

FIG. 2 Expression of *mec-2-lacZ*. a, Wild type. Plasmid TU #194 was used in all panels except b. We observed a similar axonal staining in transgenic animals carrying *lacZ* fused after the hydrophobic domain (TU #195, data not shown); b, plasmid TU #199 in wild type; c, *unc-86(e1416)*: touch-cell expression is abolished; some cells still stain in the head and tail; d, *mec-3(u6)*. Expression of *mec-2-lacZ* in the ALML/R, AVM, and PVM touch cells was greatly decreased (<5% of the cells stained); 67% of the PLML/R cells still stained ($n = 50$); e, *mec-12(u63)*; f, *mec-12(u241)*.

METHODS. Construction of the fusion plasmids is described in Fig. 1 legend. Germline transformation and X-gal staining were performed as described^{10,31}. Touch-cell axons were examined in *mec-12(u63)* animals using anti-*mec-7* antibody as described in ref. 33, and in *mec-12(u241)* animals using a *mec-7* promoter-directed GFP fluorescence¹¹. Process length in both strains was wild-type.



the outside of the microtubule bundle and usually (80%) contact the axonal membrane via diffusely staining material²⁴. Proteins linking channels with cytoskeletal proteins have also been found by others^{9,25,26}.

The linkage of the degenerin channels and the large-diameter microtubules could localize the channels, as the binding of the rat ENaC proteins to spectrin is thought to⁹. In addition, this linkage could be the basis of mechanosensation by these cells. Specifically, displacement of the microtubules could lead to the opening of the channels using MEC-2 as an intermediary. As the microtubules appear to be crosslinked to each other²⁴, the displacement of one could potentially affect the activity of several channels. We envisage that the degenerin channel is also attached to the extracellular matrix. Within the extracellular region of the *C. elegans* proteins is a domain that appears to regulate degenerin function²⁷, and attachment to the cuticle by extracellular material is important for touch-cell function¹. Movement of the channel between its cytoskeletal and extracellular attachment points would force the channel open. This model, like that of hair-cell function (reviewed in ref. 28), proposes that the physical manipulation of tethered channels underlies mechanosensory transduction. □

Received 18 May; accepted 18 September 1995.

1. Chalfie, M. & Sulston, J. *Dev. Biol.* **82**, 358–370 (1981).
2. Chalfie, M. & Au, M. *Science* **243**, 1027–1033 (1989).
3. Huang, M. & Chalfie, M. *Nature* **367**, 467–470 (1994).

4. Driscoll, M. & Chalfie, M. *Nature* **349**, 588–593 (1991).
5. Hong, K. & Driscoll, M. *Nature* **367**, 470–473 (1994).
6. Stewart, G. W. et al. *Blood* **79**, 1593–1601 (1992).
7. Stewart, G. W., Argent, A. C. & Dash B. C. *J. Biochim. biophys. Acta* **1225**, 15–25 (1993).
8. Pawson, T. & Schlessinger, J. *Curr. Biol.* **3**, 434–442 (1993).
9. Rotin, D. et al. *EMBO J.* **13**, 4440–4450 (1994).
10. Fire, A., Harrison, S. W. & Dixon, D. *Gene* **93**, 189–198 (1990).
11. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. & Prasher, D. *Science* **263**, 802–805 (1994).
12. Lewis, J. A. & Hodgkin, J. A. *J. comp. Neurol.* **172**, 489–510 (1977).
13. Bargmann, C. I., Hartwig, E. & Horvitz, H. R. *Cell* **74**, 515–527 (1993).
14. Finney, M. & Ruvkun, G. *Cell* **63**, 895–905 (1990).
15. Way, J. C. & Chalfie, M. *Genes Dev.* **3**, 1823–1833 (1989).
16. Manoil, C. *J. Bact.* **172**, 1035–1042 (1990).
17. Salzer, U., Ahorn, H. & Prohaska, R. *Biochim. biophys. Acta* **1151**, 149–152 (1993).
18. Garnier, J. *J. molec. Biol.* **120**, 97–120 (1978).
19. Chalfie, M. & Thomson, J. N. *J. Cell Biol.* **93**, 15–23 (1982).
20. Savage, C. et al. *Genes Dev.* **3**, 870–881 (1989).
21. Savage, C. et al. *J. Cell Sci.* **107**, 2165–2175 (1994).
22. Maccioni, R., Rivas, C. & Vera, J. *EMBO J.* **7**, 1957–1963 (1988).
23. Rivas, C., Vera, J., & Maccioni, R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6092–6096 (1988).
24. Chalfie, M. & Thomson, J. N. *J. Cell Biol.* **82**, 278–289 (1979).
25. Srinivasan, Y., Elmer, L., Davis, J., Bennett, V. & Angelides, K. *Nature* **333**, 177–180 (1988).
26. Kirsch, J., Wolters, I., Triller, A. & Betz, H. *Nature* **366**, 745–748 (1993).
27. García-Añoveros, J., Ma, C. & Chalfie, M. *Curr. Biol.* **5**, 441–448 (1995).
28. Hudspeth, A. J. & Gillespie, G. *Neuron* **12**, 1–9 (1994).
29. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1989).
30. Coulson, A., Sulston, J., Brenner, S. & Karn, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7821–7825 (1986).
31. Mello, C. C. et al. *EMBO J.* **10**, 3959–3970 (1991).
32. Okkema, P. & Fire, A. *Development* **120**, 2175–2186 (1994).
33. Mitani, S. et al. *Development* **119**, 773–783 (1993).

ACKNOWLEDGEMENTS. We thank M. Driscoll and A. Chambers for help with the identification of the u299 insertion site; A. Coulson and J. Sulston for cosmid; P. Okkema for the cDNA library; A. Fire for lacZ vectors; and M. Hamelin, M. Chou, J. Culotti, T. Fukushige and S. Siddiqui for sharing unpublished data on *mec-12*. Some strains were received from the *Caenorhabditis* Genetics Center, which is funded by the NIH National Center for Research Resources. This work was supported by the US NIH.

Partial activation of CD8⁺ T cells by a self-derived peptide

Wuxiong Cao^{*†}, Scott S. Tykodi^{*†},
Mark T. Esser^{*†}, Vivian L. Braciale^{*†}
& Thomas J. Braciale^{*†§}

^{*} Beirne B. Carter Center for Immunology Research, and the Departments of [†] Microbiology and [§] Pathology, University of Virginia Health Sciences Center, Charlottesville, Virginia 22908, USA

[‡] Division of Biology and Biomedical Science, Immunology Program, Washington University School of Medicine, St Louis, Missouri 63110, USA

T CELLS are normally activated when the peptide for which they are specific is presented to them in the context of the appropriate major histocompatibility complex (MHC) (class I and Class II for CD8⁺ and CD4⁺ T cells, respectively). An increasing body of evidence indicates that structural homologues of the immunogenic peptide can partially activate or antagonize CD4⁺ T cells^{1–3}. CD8⁺ T cells may also be partially antagonized by such peptides^{4,5}, and self-derived peptides of this type may play a role in CD8⁺ T cell selection in the thymus^{6–8}. Activated CD8⁺ T cells lyse their targets by perforin-dependent granule exocytosis^{9,10} and by inducing apoptosis mediated by CD95 (also known as Fas or APO1) with its ligand (CD95L)^{11–15}. Here we show that a clone of K^d-restricted CD8⁺ T cells specific for influenza haemagglutinin, which can also be activated in a crossreactive manner by a peptide derived from a myeloma tumour immunoglobulin heavy-chain variable region (IgVH) to kill by both routes¹⁶, kills only by the CD95-CD95L pathway when stimulated by the corresponding germline IgVH peptide. As this germline IgVH peptide differs from the tumour peptide only at a single position buried in the MHC-binding groove, this indicates that CD95-CD95L-mediated killing can be triggered independently of the perforin-mediated pathway, and can be selectively affected by changes in MHC conformation.

We have previously reported that an H-2K^d-restricted cytotoxic T-lymphocyte (CTL) clone, 40-2, specific for a haemagglutinin (HA) epitope (JHA 210–219; Fig. 1a) of influenza virus strain A/Japan/57 (H2N2) can also recognize an endogenously processed IgVH sequence, IgVH residues 49–58 of the myeloma tumour MOPC 21 in association with H-2K^d (ref. 16). The corresponding germline MOPC 21 VH gene differs from the myeloma VH at residue 58, where isoleucine replaces leucine in the myeloma VH sequence (Fig. 1a and W.C. et al., manuscript in preparation). As residue 58 is the carboxy-terminal residue of the minimal MOPC 21 decamer peptide VH 49–58 recognized by the HA-specific CTL clone 40-2 (ref. 16), we investigated whether this conservative substitution in the C-terminal K^d-binding anchor residue could affect the binding and crossreactive recognition of the germline IgVH sequence by clone 40-2. As Fig. 1b shows, synthetic peptides corresponding to either the myeloma epitope 21VH-L or the germline sequence GL VH-I compete for K^d binding with a similar efficiency¹⁷, suggesting that K^d binding is comparable. This conclusion was confirmed in an MHC stabilization assay using RMA-S cells expressing the transfected K^d molecule (Fig. 1c).

Unlike the tumour peptide, the germline GL VH-I peptide failed to sensitize P815 cells to CTL-mediated lysis over a range of peptide concentrations in ⁵¹Cr-release cytotoxicity assays carried out for up to 12 hours (Fig. 1d). The isoleucine-for-leucine substitution in the ten-amino-acid VH peptide could not easily account for this non-reactivity as an isoleucine-substituted HA210–219 decamer peptide (JHA210–219-I) (Fig. 1a) was recognized more efficiently than the corresponding wild-type viral peptide (Fig. 1d). Clone 40-2 maintained its specificity and did not recognize target cells displaying the corresponding 210–219 epitope of the A/AA/57 HA (AHA 210–219), which has a glycine-to-serine substitution at position 215 (ref. 16; Fig. 1a, d).

A variant has recently been identified of an unrelated CD8⁺ CTL clone that has lost the capacity to trigger antigen-receptor-mediated cytotoxicity through the perforin pathway but can still provoke antigen-receptor-dependent CD95-mediated cytotoxicity (M.T.E. et al., manuscript submitted). This mutant CTL clone triggers CD95/CD95L-dependent killing¹⁸ of target cells, as measured by DNA fragmentation before disintegration of the

§ To whom correspondence should be addressed.

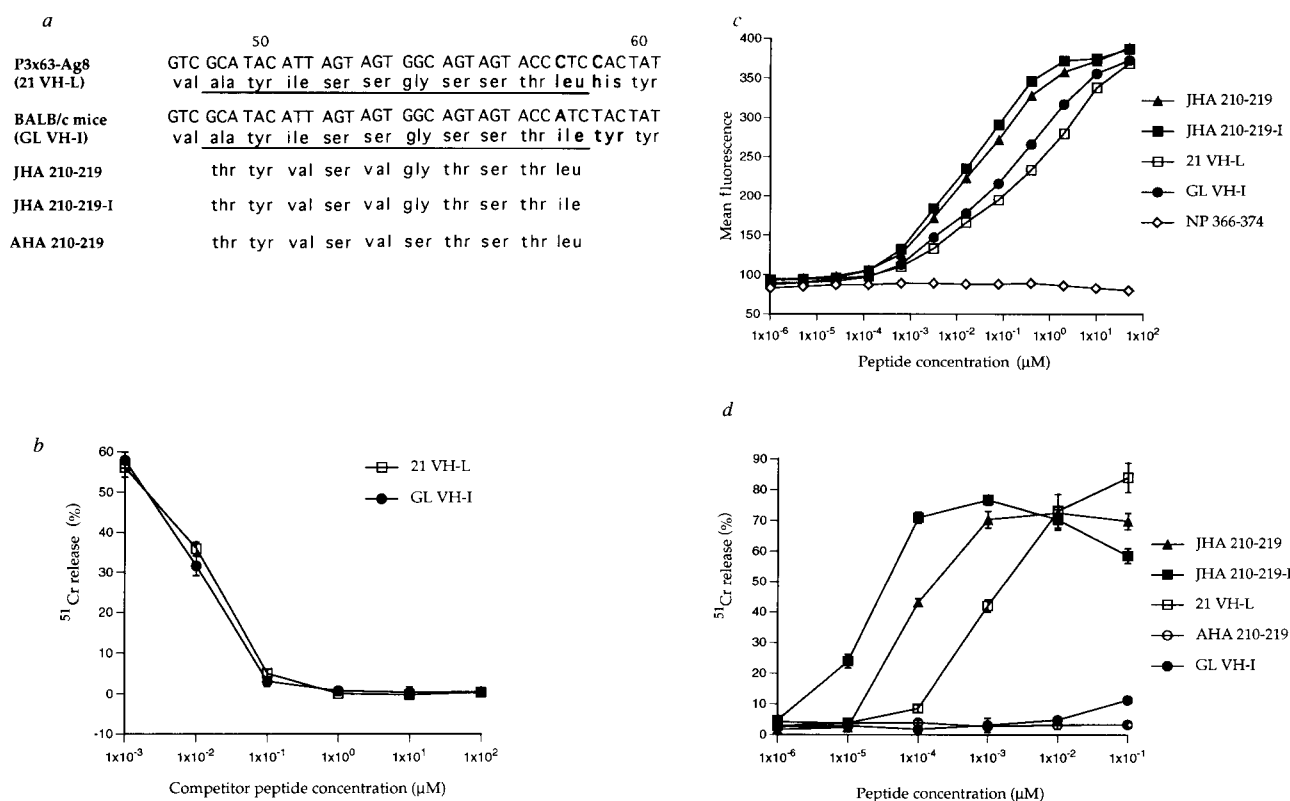


FIG. 1 **a**, Derived sequences of MOPC 21 myeloma and mouse germline IgVH segments. Nucleotide and amino-acid differences are indicated by bold letters. The JHA210–219-homologous regions are underlined and the corresponding HA210–219 sequences are also shown. **b**, Inhibition of CTL-mediated lysis of NP147–155-sensitized targets by MOPC 21 VH-L and GL VH-I peptides. Peptide competition cytotoxicity assay was carried out as described¹⁶. **c**, Stabilization of K^d molecules on the surface of RMA-S cells transfected with K^d. RMA-S cells transfected with

the K^d gene were incubated with synthetic peptides for 4 h at room temperature and 7 h at 37 °C. Flow-cytometry analysis was used to measure K^d on the cell surface. An H-2D^b-binding peptide NP366–374 did not stabilize K^d on the cell surface. **d**, Recognition of peptide-treated P815 target cells by CTL clone 40-2 in a 12-h standard ⁵¹Cr-release assay¹⁶. An effector-to-target ratio of 5 was used for all assays. Nonspecific killing was 4.5%. Error bars indicate standard deviations.

plasma membrane and ⁵¹Cr release. To test whether this self VH peptide could signal clone 40-2 selectively to trigger perforin-independent cytotoxicity, ³H-thymidine-labelled P815 cells were used as targets in a DNA-fragmentation-based CTL cytotoxicity assay¹⁹. As shown in Fig. 2, P815 cells treated with the germline GL VH-I peptide were recognized by clone 40-2 as detected by DNA fragmentation. As observed in the ⁵¹Cr release assay (Fig. 1d), the isoleucine-substituted JHA210–219-I peptide, but not AHA210–219 peptide, also triggered DNA fragmentation by clone 40-2. An unrelated JHA210–219-reactive CTL clone NI-12, which crossreacts on the AHA210–219 epitope but not the MOPC 21 VH-L tumour peptide in the ⁵¹Cr-release assay¹⁶, gives an identical recognition pattern by DNA fragmentation (Fig. 2). These results indicate that the difference in CTL recognition of germline VH peptide as detected in the ⁵¹Cr-release and ³H-

thymidine-release cytotoxicity assays may not reflect a greater sensitivity or a difference in specificity of the DNA-fragmentation cytotoxicity assay.

To determine whether DNA fragmentation triggered by germline VH peptide is CD95-dependent, we tested the ability of clone 40-2 to trigger ³H-thymidine release on H-2K^d L1210 cells, which either overexpress the murine *fas* gene product CD95 by transfection (L1210F⁺)²⁰ or lack detectable CD95 antigen (L1210F⁻) as a result of transfection of the cells with an antisense version of the murine *fas* gene¹⁸. As shown in Fig. 3a, germline VH-peptide triggers the cytotoxicity by clone 40-2 only on peptide-treated L1210F⁺ targets, whereas JHA210–219 and MOPC 21 VH peptides trigger cytotoxicity on both targets. L1210F⁺ target cells either treated with irrelevant AHA210–219 peptide or left untreated were not lysed by clone 40-2. The NI-

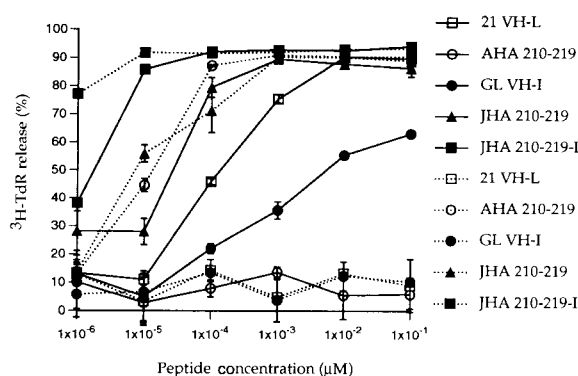


FIG. 2 Recognition of peptide-treated P815 target cells by CTL clones in the DNA-fragmentation assay. The DNA fragmentation was assayed as described¹⁹ with modification: [³H]thymidine([³H]TdR)-labelled P815 target cells were incubated alone or in the presence of synthetic peptides with 40-2 (solid lines) or NI 12 (dotted lines) cells for 12 h at 37 °C, in parallel with a ⁵¹Cr-release assay (Fig. 1d). Per cent [³H]TdR release from triplicate samples was calculated as before¹⁹. Nonspecific release was 2.9% by clone 40-2, and 7.1% by clone NI 12.

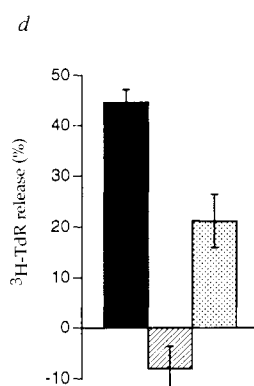
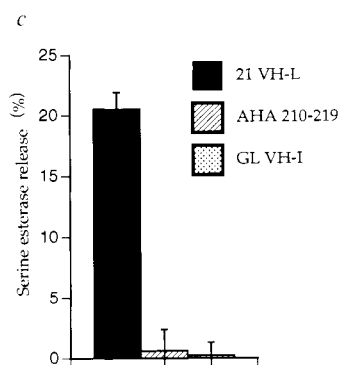
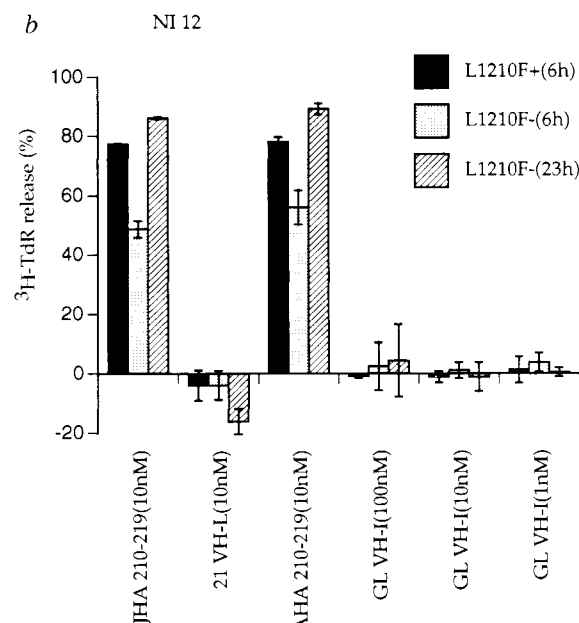
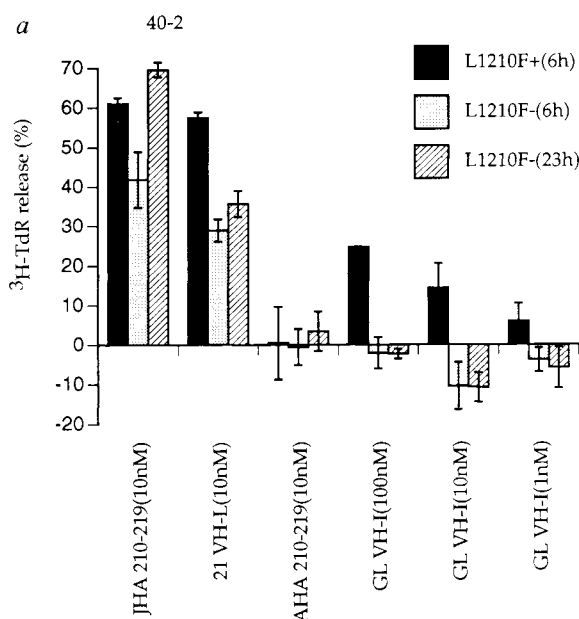


FIG. 3 GL VH-I peptide activates only the CD95 lytic pathway. *a*, GL VH-I peptide recognition by clone 40-2 on L1210F⁺ and L1210F⁻ target cells. *b*, GL VH-I peptide recognition by clone NI 12 on L1210F⁺ and L1210F⁻ target cells. Synthetic peptides at the concentrations shown were used in 6-h or 23-h DNA-fragmentation assays. Background release from non-peptide-treated L1210F⁻ cells was 32% by clone 40-2 and 14.6% by clone NI 12. Background release from L1210F⁻ cells was zero by both 40-2 and NI 12 in a 6-h assay, 13.5% by 40-2 and zero by NI 12 in a 23-h assay. Background release has been subtracted. *c* and *d*, GL VH-I peptide cannot induce serine esterase release. The serine esterase release assay (*c*) and DNA-fragmentation assay (*d*) were done in parallel using L1210F⁺ target cells. For both assays, equivalent number of CTLs, specific peptides (10 nM) and unlabelled target cells were used^{19,21}. After 6 h of incubation at 37 °C, 20 µl supernatant was removed and assayed for serine esterase as described²¹ except that the assay was carried out at pH 8.2. The percentage specific serine esterase release was calculated as the percent specific ⁵¹Cr release (Fig. 1*d*). Background release of 21.8% in the DNA-fragmentation assay has been subtracted.

12 clone, in contrast, did not recognize L1210F⁺ treated with either the MOPC 21 VH-L or GL VH-I peptide, but did recognize both L1210F⁺ and L1210F⁻ cells treated with JHA210-219 or AHA210-219 viral peptides (Fig. 3*b*).

Perforin-mediated Ca²⁺-dependent lysis¹⁸, but not Ca²⁺-independent lysis, is associated with the release of serine esterase from CTLs in preformed granules^{20,21}. Figure 3*c* and *d* compares esterase release and DNA fragmentation by clone 40-2 and shows that CTL-mediated DNA fragmentation of CD95⁺ L1210 target cells treated with tumour VH peptide but not germline VH peptide is associated with serine esterase release by clone 40-2. This result supports the view that cytolysis triggered by

germline VH-I peptide does not trigger the perforin/serine esterase lytic mechanism.

CD95-mediated lysis requires protein-synthesis-dependent¹⁸ CD95-ligand induction in CTL^{22,23}. As shown in Fig. 4, cytolysis by clone 40-2 of CD95⁺ target cells treated with germline VH peptide is blocked by the protein-synthesis inhibitor emetine. Emetine had only a modest effect on 40-2-mediated cytolysis of target cells treated with the MOPC 21 VH peptide, probably reflecting CD95-ligand-independent DNA fragmentation mediated by preformed perforin granules.

In conclusion, our observations indicate that engagement of the T-cell antigen receptor (TCR) on CD8⁺ CTL by a self pep-

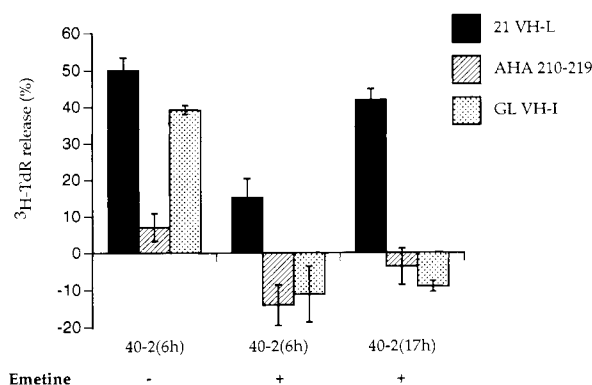


FIG. 4 Protein synthesis is required for activation of the CD95 pathway by GL VH-I. [³H]TdR-labelled L1210F⁺ cells were incubated with synthetic peptides (10 nM) and clone 40-2 in the absence or presence of 0.6 µg ml⁻¹ emetine for 6 or 17 h. Background release by 40-2 has been subtracted and was 21.3% in the 6-h assay without emetine, and with emetine was 2.7% in the 6-h assay and 23.5% in the 17-h assay.

tide can selectively trigger CD95-dependent cytotoxicity, implying that the signalling pathways for induction of perforin- and CD95-dependent cytotoxicity may be different, as suggested by recent observations (M.T.E. *et al.*, manuscript submitted). Furthermore, as the self germline VH peptide differs from the non-self MOPC 21 sequence by a single methyl group on the side chain of a dominant hydrophobic anchor residue for K^d, our finding implies that changes either in TCR-accessible residues²⁴ or in residues buried in the MHC-binding groove can alter the conformation of the peptide-MHC complex sufficiently to result in partial activation^{25,26}. The signal transduced by the self pep-

tide leading to selective activation of the CD95 cytotoxicity pathway in this activated CTL population may even mimic a partial signalling event triggered by self peptides in resting CD8⁺ T cells *in vivo*. The CD8⁺ T-cell response to the HA210-219 epitope is notably weak and variable in H-2^d haplotype mice owing to a low frequency of HA210-219 CTL precursors (W.C. *et al.*, manuscript in preparation). Considering the evidence implicating the CD95 pathway in the regulation of T-lymphocyte responses²⁷⁻²⁹, the weak HA210-219 response may reflect partial antagonism of CD8⁺ T-cell precursors *in vivo* by this endogenously processed IgVH self peptide. □

Received 25 April; accepted 29 August 1995.

1. Evavold, B. D. & Allen, P. M. *Science* **252**, 1308-1310 (1991).
2. Evavold, B. D., Williams, S. G., Hsu, B. L., Buus, S. & Allen, P. M. *J. Immunol.* **148**, 347-353 (1992).
3. Sette, A. *et al.* *A. Rev. Immunol.* **12**, 413-431 (1994).
4. Jameson, S. C., Carbone, F. R. & Bevan, M. J. *J. exp. Med.* **177**, 1541-1550 (1993).
5. Bertoletti, A. *et al.* *Nature* **369**, 407-410 (1994).
6. Jameson, S. C., Hogquist, K. A. & Bevan, M. J. *Nature* **369**, 750-752 (1994).
7. Spain, L. M., Jorgensen, J. L., Davis, M. M. & Berg, L. J. *J. Immunol.* **152**, 1709-1717 (1994).
8. Hogquist, K. A. *et al.* *Cell* **76**, 17-27 (1994).
9. Henkart, P. A. *A. Rev. Immunol.* **3**, 31-58 (1985).
10. Kagi, D. *et al.* *Nature* **369**, 31-37 (1994).
11. Ito, N. *et al.* *Cell* **66**, 233-243 (1991).
12. Ogasawara, J. *et al.* *Nature* **364**, 806-809 (1993).
13. Lowin, B., Hahne, M., Mattmann, C. & Tschopp, J. *Nature* **370**, 650-652 (1994).
14. Stalder, T., Hahn, S. & Erb, P. *J. Immunol.* **152**, 1127-1133 (1994).
15. Kagi, D. *et al.* *Science* **265**, 528-530 (1994).
16. Cao, W., Myers-Powell, B. A. & Braciale, T. J. *J. exp. Med.* **179**, 195-202 (1994).

17. Bodmer, H. C. *et al.* *Cell* **52**, 253-258 (1988).
18. Walsh, C. M., Glass, A. A., Chiu, V. & Clark, W. R. *J. Immunol.* **153**, 2506-2514 (1994).
19. Matzinger, P. *J. Immunol. Meth.* **145**, 185-192 (1991).
20. Rouvier, E., Luciani, M.-F. & Golstein, P. *J. exp. Med.* **177**, 195-200 (1993).
21. Ostergaard, H. L., Kane, K. P., Mescher, M. F. & Clark, W. R. *Nature* **330**, 71-72 (1987).
22. Suda, T., Takahashi, T., Golstein, P. & Nagata, S. *Cell* **75**, 1169-1178 (1993).
23. Suda, T. & Nagata, S. *J. exp. Med.* **179**, 873-879 (1994).
24. Rammensee, H. G., Falk, K. & Rotzschke, O. *A. Rev. Immunol.* **11**, 213-244 (1993).
25. Sherman, L. A. & Chattopadhyay, S. *A. Rev. Immunol.* **11**, 385-402 (1993).
26. Chattopadhyay, S., Theohab, M., Biggs, J. & Sherman, L. A. *J. exp. Med.* **179**, 213-219 (1993).
27. Henkart, P. A. *Immunity* **1**, 343-346 (1994).
28. Crispe, I. N. *Immunity* **1**, 347-349 (1994).
29. Singer, G. G. & Abbas, A. K. *Immunity* **1**, 365-371 (1994).

ACKNOWLEDGEMENTS. We thank W. R. Clark for CD95-positive and -negative L1210 cells, M. Bevan for the RMA-S cell line transfected with K^d, and A. Srikiachorn for modifications to the DNA-fragmentation assay. This work is supported by grants from the USPHS. The continuing support of the Beirne B. Carter Foundation is gratefully acknowledged.

Perinatal lethality and blocked B-cell development in mice lacking the tyrosine kinase Syk

Martin Turner, P. Joseph Mee, Patrick S. Costello, Owen Williams, Abigail A. Price, Linda P. Duddy, Michael T. Furlong*, Robert L. Geahlen* & Victor L. J. Tybulewicz†

National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

* Department of Medicinal Chemistry and Pharmacognosy, Purdue University, West Lafayette, Indiana 47907, USA

THE tyrosine kinase Syk (relative molecular mass 72,000), which is widely expressed in haematopoietic cells, becomes associated with and activated by engagement of the B-cell antigen receptor^{1,2}. Furthermore, it has been implicated in signalling through the receptors for interleukin-2 (IL-2)³, granulocyte colony-stimulating factor (G-CSF)⁴ and Fc ϵ , the T cell receptor, as well as through receptors for several platelet agonists⁵. A homologous kinase, ZAP-70, is crucial in signalling through the T-cell receptor and in T-cell development^{6,9}. Using homologous recombination in embryonic stem cells, we created mice null for the *syk* gene which showed petechiae *in utero* and died shortly after birth. Irradiated mice reconstituted with Syk-deficient fetal liver showed a block in B-cell development at the pro-B to pre-B cell transition, consistent with a key role for Syk in pre-B-cell receptor signalling. Despite the production of small numbers of immature B cells, Syk-deficient radiation chimaeras failed to accumulate mature B cells, indicating a possible role for this protein in the production or maintenance of mature B cells. In addition, whereas the development of $\alpha\beta$ T cells proceeded normally, Syk-deficient mice showed impaired development of thymocytes using the V γ 3 variable region gene (V γ 3⁻ thymocytes). Finally, we show that Syk is not required for signalling through the IL-2 and G-CSF receptors.

A mutation was introduced into the kinase domain of the *syk* gene by homologous recombination in embryonic stem (ES) cells (Fig. 1a). Two such targeted clones were used to establish mouse strains carrying the mutation and mice heterozygous for the mutation (*syk*^{+/−}) were intercrossed to generate homozygous mutants (*syk*^{−/−}) (Fig. 1b). Immunoblotting experiments showed no detectable Syk, confirming the inactivation of the *syk* gene (Fig. 1c).

Genotype analysis of embryos from heterozygous intercrosses at days 16.5 and 18.5 of gestation (E16.5 and E18.5) showed the expected numbers of *syk*^{−/−} fetuses (Table 1). A large number of pups died shortly after birth, however, and these were almost always *syk*^{−/−}. A deficit of mutants was apparent 2 days after birth, and of 227 mice analysed at 10–20 days of age, only 1 was *syk*^{−/−}.

At E16.5, all *syk*^{−/−} fetuses showed extensive haemorrhaging (Fig. 1d, Table 1). Interestingly, this phenotype seems to be transient, as at E18.5 it occurred in only 45%. The petechiae are unlikely to be due to a platelet deficiency or dysfunction, because mice lacking platelets do not show *in utero* haemorrhaging¹⁰. Rather, it is more likely to be caused by a defect in vascular

TABLE 1 Genotype and phenotype of offspring derived from heterozygous parents

Age	No. of mice analysed (litters)	Genotype			No. of mutants identified visually (% of all mutants)
		+/+	+/-	-/-	
E16.5	125 (13)	27	61	37	37 (100)
E18.5	90 (9)	21	47	22	10 (45)
2 days	65 (6)	20	36	9	9 (100)
>10 days	227 (19)	71	155	1	0 (0)

Mice heterozygous for the *syk* mutation were intercrossed. Progeny were analysed either at 16.5 or 18.5 days of gestation (E16.5, E18.5) or at 2 or >10 days after birth for genotype by Southern blot analysis of fetal liver or tail DNA. Genotypes were identified as wild-type (+/+), heterozygous (+/-) or homozygous (-/-) mutant. The number and percentage of mutant mice showing a visible phenotype are indicated. Fetuses were scored for petechiae and neonatal pups for chylous ascites and/or runting and failure to thrive.

† To whom correspondence should be addressed.

tissue. Similar haemorrhaging was observed in mice deficient in platelet-derived growth factor (PDGF)-B¹¹ and PDGF receptor- β ¹², where the defect was postulated to be caused by a failure of communication between platelets and endothelial cells or between endothelial cells and smooth muscle cells.

Once *syk*^{-/-} pups had fed, they often accumulated large quantities of chylous ascites containing many lipid-laden macrophages and chylomicrons (Fig. 1e). A naturally occurring mouse mutation, *chy*, has been reported with chylous ascites, and it was suggested that the phenotype could have resulted from defective

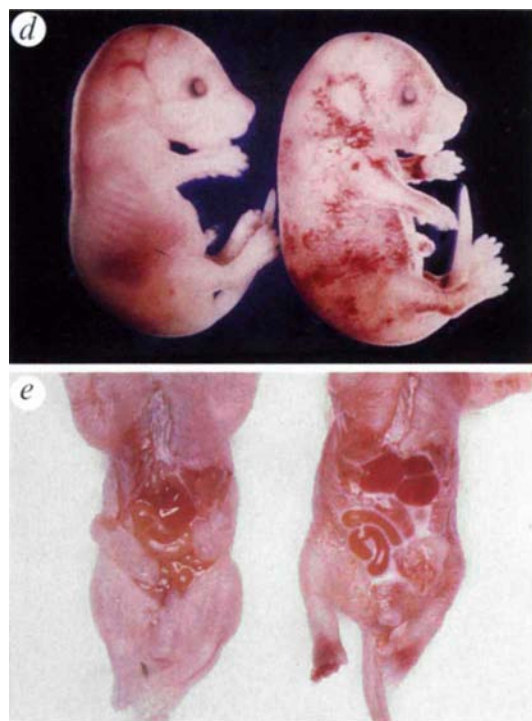
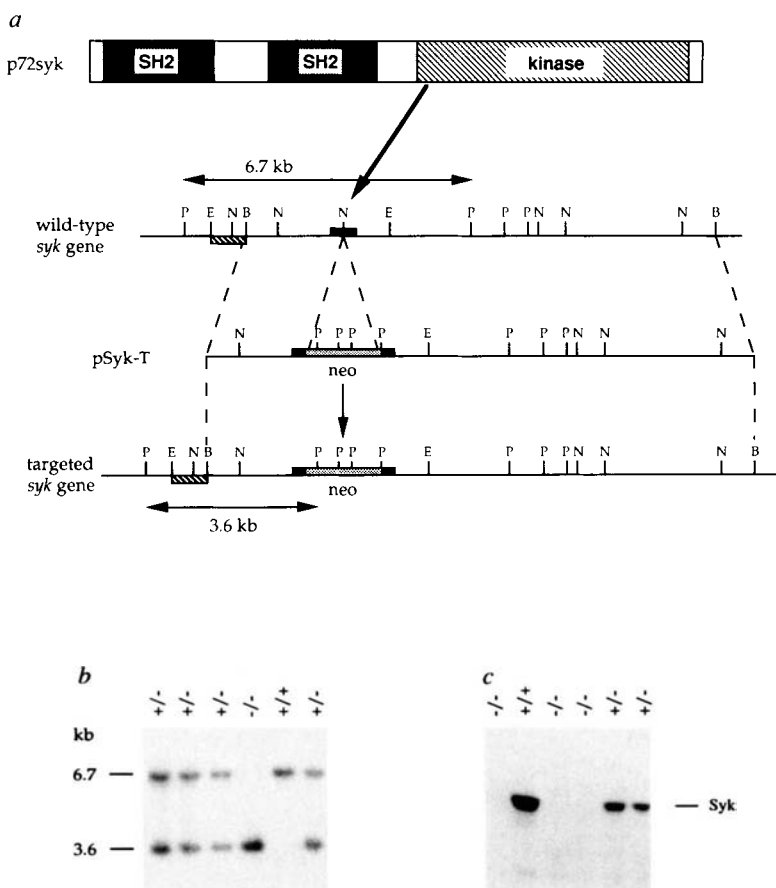


FIG. 1 Targeted disruption of the *syk* gene. **a**, The *syk* gene was disrupted at an *NcoI* site located in an exon containing sub-domain I of the kinase domain²⁶. The targeting vector pSyk-T and the predicted structures of the wild-type and targeted *syk* genes are indicated. The top line shows the location of the mutation on a diagram of the structure of *Syk*. On the gene maps, the black and shaded boxes represent the targeted exon and the neomycin-resistance gene, respectively. The sizes of the two predicted *PstI* fragments hybridizing to the diagnostic probe (hatched box) are shown. Restriction enzyme sites: B, *Bam*HI; E, *Eco*RI; N, *Nco*I; P, *Pst*I. **b**, Southern blot analysis of tail DNA from 2-week-old mice. *PstI*-digested DNA was hybridized with the probe shown above. Mice were identified as wild-type (+/+), heterozygous mutant (+/-) or homozygous mutant (-/-). **c**, Immunoblot analysis of *Syk*. Immunoblots of lysates of fetal livers were stained with an anti-*Syk* rabbit antiserum raised against the peptide EPTGGPWGPDGRL corresponding to amino acids 318–330 in murine *Syk* (M.T.F., K. H. Kim and R.L.G., GenBank accession number U25685). The absence of *Syk* was confirmed with immunoblots using a rabbit antiserum raised against a peptide from mouse *Syk* corresponding to amino acids 271–289 of murine *Syk*, a serum against the two SH2 domains of mouse *Syk* and a serum against the kinase domain of *Syk* (data not shown). Each of these four sera failed to detect any full-length or truncated forms of *Syk*. **d**, Photograph of a wild-type (left) and *syk*^{-/-} (right) fetus at 16.5 days of gestation. The mutant shows petechiae. We observed this phenotype with *syk* mice derived from both of the two independent *syk*^{-/-} ES cell lines. In contrast, *syk*^{+/-} mice were always indistinguishable from wild-type mice. **e**, Photograph of a wild-type (left) and *syk*^{-/-} (right) 2-day-old pup dissected to show chylous ascites in the peritoneal cavity of the mutant. Note also the haemorrhaging apparent in the hind feet of the mutant pup.

METHODS. λ phage containing about 30 kilobases (kb) of *syk* genomic

DNA were isolated from a genomic library of 129/Sv mouse DNA in λ FIXII (Stratagene) using a fragment of mouse *syk* complementary DNA (nucleotides 479–1,182) (M.T.F., K. H. Kim and R.L.G., GenBank accession number U25685). An 11-kb *Bam*HI fragment containing exons encoding part of the kinase domain of *Syk* was subcloned into pSC3Z (a gift from R. Gregory, Genzyme), a low-copy-number plasmid that permitted its amplification without rearrangement, to generate pSC3Z-*Syk*11B. A 1.8-kb *Xho*I-*Xba*I fragment from pPNT²⁷ containing PGKneo was blunted and inserted into a blunted *Nco*I site in pSC3Z-*Syk*11B (nucleotide 1,007 in the murine *Syk* cDNA) that lay within an exon containing sub-domain I of the kinase domain, to generate pSyk-T. The structure of the plasmid was confirmed by sequence analysis across all ligated joins. The *neo* gene in pSyk-T was in the opposite transcriptional orientation to the *syk* gene. pSyk-T was cut with *Bam*HI to release the targeting construct with ends that were homologous to the *syk* gene. This DNA was electrophorated into D3 ES cells and G418-resistant ES lines were isolated as previously described²⁷. These lines were screened for the correct targeting event by Southern blot analysis of *Pst*I-cut DNA hybridized with a 400-bp *Eco*RI/*Bam*HI fragment located immediately 5' of the targeting DNA. We identified 7 targeted colonies in 375 colonies screened. The correct structure of the targeted *syk* gene was confirmed by further digests with *Bam*HI, *Eco*RI and *Sac*I and hybridization with the 5' probe described above, with a DNA fragment located 3' to the targeting DNA and with the *neo* gene (data not shown). Targeted ES cells were injected into C57Bl/6 blastocysts as described²⁷ and resulting chimaeric mice were bred against (B6D2)_F₁ mice to transmit the mutation. Two independent targeted ES lines (*Syk*40 and *Syk*328) transmitted the mutation. Homozygous mutant mice were generated by interbreeding of heterozygous mutants. Immunoblotting was carried out using standard techniques on cell lysates of E18.5 fetal livers. Antibody binding was revealed with ECL (Amersham).

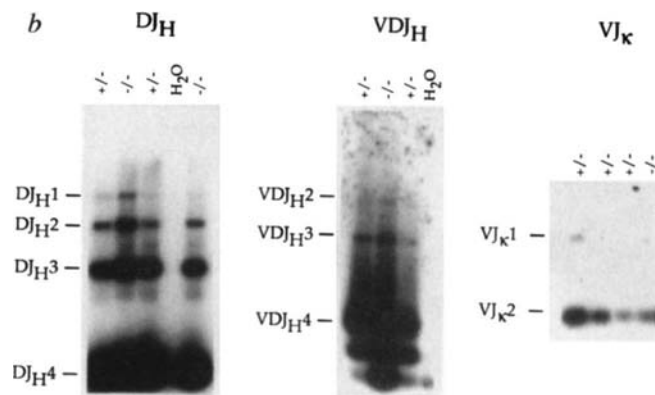
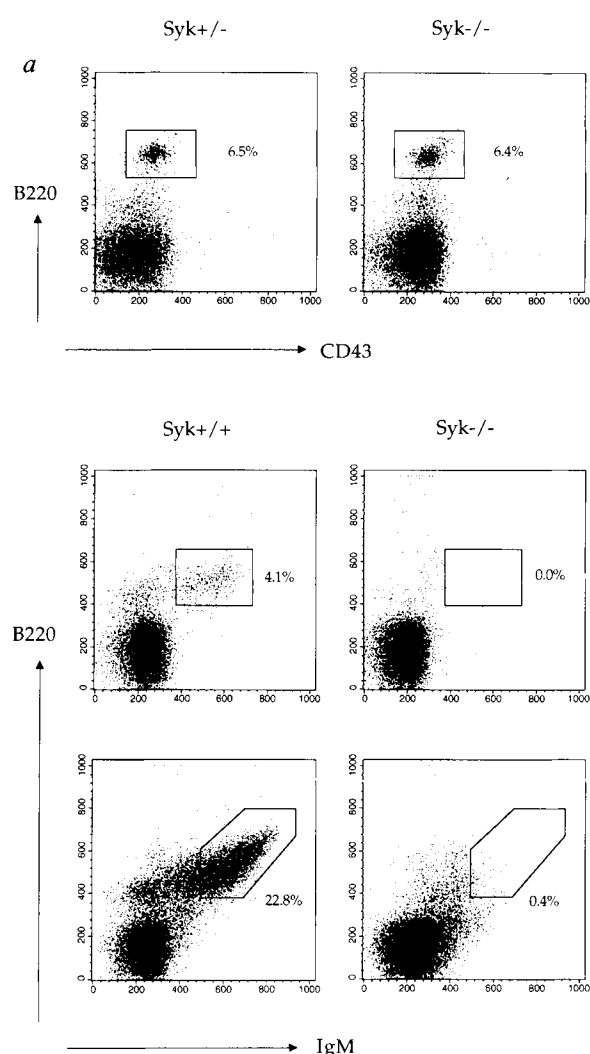


FIG. 2 B-cell development in *syk*^{-/-} mice. **a**, Flow cytometric analysis of staining of E18.5 fetal liver cells is shown in the top panel; the box identifies B-cell progenitors (B220⁺CD43⁺)¹⁴. Staining of splenocytes from 2-day-old pups is shown in the middle panel and that of E16.5 fetal liver cells, cultured on ST2 stromal cells for 10 days to generate B cells, is shown in the lower panel. In these cases the boxes identify B cells (B220⁺IgM⁺). Also shown are percentages of cells that lie in the boxes relative to all cells falling within the lymphocyte scatter gate. **b**, Polymerase chain reaction analysis of diversity-joining (DJ) and variable-DJ (VDJ) joints in the immunoglobulin heavy-chain genes and of VJ joints in the κ light-chain genes carried out on DNA from E18.5 fetal livers. Genotypes of fetuses are indicated. In some cases water was added instead of DNA as a negative control. All negative controls always failed to amplify any fragments. **c**, Flow cytometric analysis of bone marrow from B6 radiation chimaeras reconstituted with *syk*^{+/+} or *syk*^{-/-} fetal liver and from a RAG-1-deficient mouse. In all cases, staining with Ly9.1 showed that all B220⁺ cells in the chimaeras were donor in origin (data not shown). The upper panels have three regions drawn on them corresponding to pro- and pre-B cells (B220⁺IgM⁺), immature B cells (B220⁺IgM⁺) and mature B cells (B220⁺IgM⁺). The *syk*^{-/-} chimaera shows very few immature and even fewer mature B cells and these are absent in the RAG-1^{-/-} mouse. The lower panels show the level of CD25 on the pro- and pre-B cells gated as in the upper panel for cells that are B220⁺IgM⁺. Horizontal lines mark CD25⁺ and CD25⁻ populations used to identify pro- and pre-B cells, respectively. Also shown are percentages of cells that lie in the boxes relative to all cells falling within the lymphocyte scatter gate. **d**, Quantification of numbers of pro-, pre-, immature and mature B cells in B6 radiation chimaeras reconstituted with *syk*^{+/+} (open columns) and *syk*^{-/-} (shaded columns) fetal liver. Cell types were counted using the gates and staining levels as defined in **c**. Columns and error bars represent the mean and s.e.m. of numbers of cells found in two femurs of the radiation chimaeras. Analysis was done on 3 *syk*^{+/+} and 4 *syk*^{-/-} chimaeras. The *syk*^{-/-} chimaeras had significantly fewer pre-B, immature B and mature B cells ($P < 0.10$, 0.01 and 0.003 , respectively). In all cases, staining with Ly9.1 demonstrated that all B220⁺ cells were donor in origin (data not shown). The diagram at the bottom represents B-cell development and shows the surface staining characteristics of the cells enumerated. The position of developmental blocks caused by mutations in *Rag-1*, λ , μ , λ 5 and *IL-7* genes is shown^{17,20-22,28}. In this study we have shown a role for Syk in the pro-B to pre-B cell transition. The question mark indicates that Syk may also be involved in the generation of mature B cells. **e**, Flow cytometric analysis of B lymphocytes from peripheral lymphoid organs of RAG-1 radiation chimaeras reconstituted with *syk*^{+/+} or *syk*^{-/-} fetal liver: spleen (top panel); lymph node (middle panel); and peritoneal cavity (lower panel).

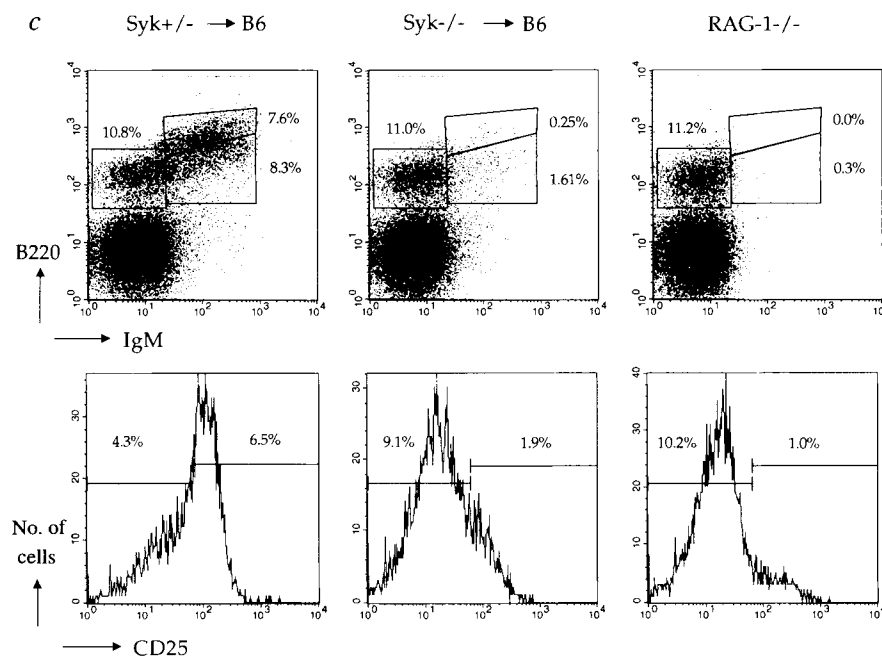


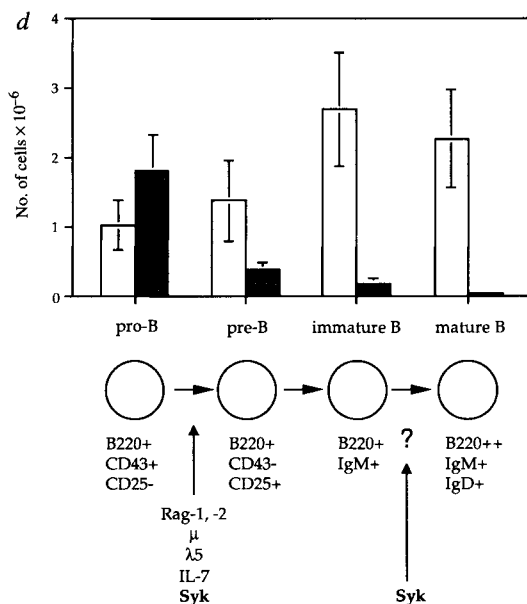
FIG. 2c Flow cytometric analysis of B lymphocytes from peripheral lymphoid organs of RAG-1 radiation chimaeras reconstituted with *syk*^{+/+} or *syk*^{-/-} fetal liver: spleen (top panel); lymph node (middle panel); and peritoneal cavity (lower panel).

METHODS. Fetal liver cells were cultured as a single-cell suspension (2×10^6 cells ml⁻¹) on a layer of confluent irradiated ST2 bone marrow stromal cells¹⁶ in RPMI-1640 containing fetal calf serum (10%), antibiotics, 2-mercaptoethanol (5×10^{-5} M) and murine IL-7 (100 U ml⁻¹; Genzyme). Cultures were refed on day 7 with the

lymphatic ducts¹³. A similar explanation may also apply in *syk*^{-/-} mice.

There was no reduction in the numbers of erythrocytes, leukocytes or platelets in the peripheral blood of E16.5 and E18.5 *syk*^{-/-} fetuses, and the numbers of E16.5 fetal liver progenitors responsive to IL-3, G-CSF, macrophage(M)-CSF and GM-CSF were unaltered in *syk*^{-/-} mutants (not shown). Thus the absence of Syk does not grossly perturb haematopoiesis. Syk has been reported to associate with the G-CSF receptor and to become activated following ligand binding⁴. Our data show that, at least in myeloid progenitors, Syk is not essential for G-CSF-receptor signalling.

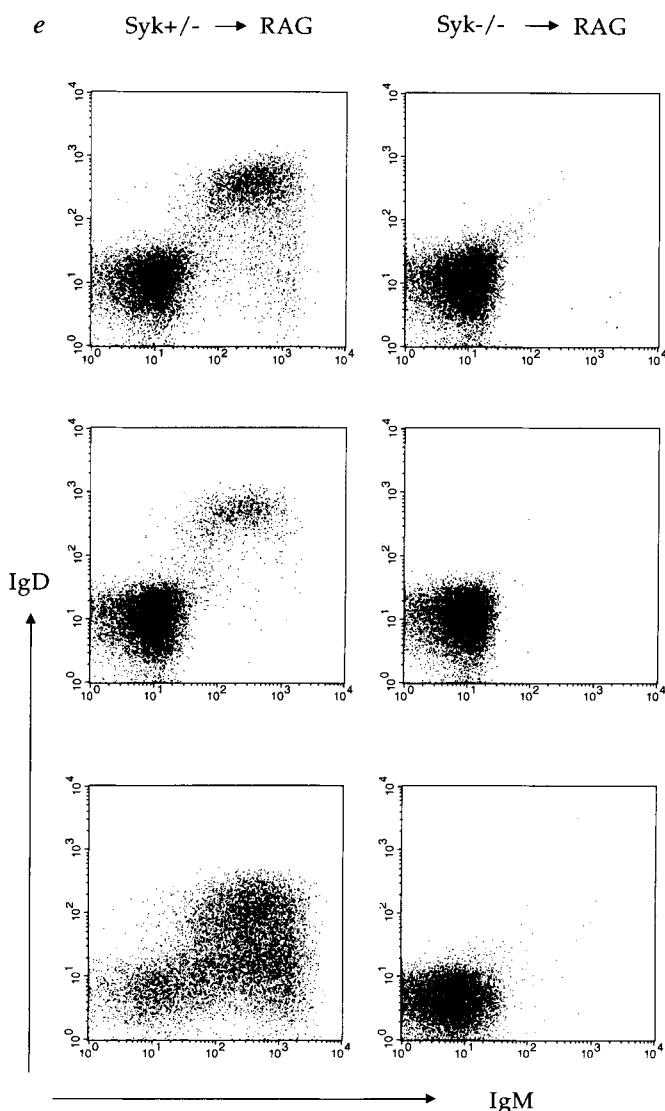
The numbers of progenitor B cells in *syk*^{-/-} embryos was unaltered as shown by flow cytometry (Fig. 2a)¹⁴ and by a B-cell soft agar colony-forming assay¹⁵ on ST2 stromal cells¹⁶ (not shown). Polymerase chain reaction (PCR) analysis of *syk*^{-/-} fetal liver DNA revealed the presence of rearrangements in both immunoglobulin heavy- and light-chain genes (Fig. 2b). Thus the early stages of B-cell development proceed normally. In contrast, the spleens of 2-day-old *syk*^{-/-} mice had no detectable IgM-bearing (IgM⁺) B cells and *syk*^{-/-} fetal liver cells cultured on ST2 cells failed to produce IgM⁺ B cells (Fig. 2a). These experiments suggested that there may be a block in B-cell development in *syk*^{-/-} mice.



► same medium lacking IL-7 and a further 3 days later non-adherent cells were removed and stained with anti-IgM-FITC (fluorescein isothiocyanate), anti-B220-allophycocyanin (APC) and propidium iodide (PI). Dead cells were excluded on the basis of low forward-light scatter and PI staining. Immunoglobulin rearrangements were assayed by PCR using previously described oligonucleotides and revealed by Southern blotting of the products followed by hybridization with probes as described previously^{29,30}. Radiation chimaeras were made by injecting 1.5×10^6 fetal liver cells intravenously into either female C57Bl/6 or RAG-1-deficient mice¹⁷ that had received 1,000 or 600 rad, respectively, from a ⁶⁰Co source. The recipient mice were given neomycin sulphate (0.16%) in their drinking water for 4 weeks following irradiation and analysed 2–5 months following reconstitution. To enumerate B-cell subsets in bone marrow, cells were flushed from two femurs, and the proportion of total cells that were lymphocytes was calculated by electronic gating by FACS. Staining with the indicated combinations of monoclonal antibodies was done in the presence of 10% heat-inactivated normal rat serum and fluorescence was analysed on a FACSstar flow cytometer with Cellquest software (Becton-Dickinson). Dead cells were excluded on the basis of low forward-light scatter and only live cells falling within the lymphocyte scatter gate are shown. Anti-B220-APC and anti-CD43-FITC were from Pharmingen, anti-IgM-FITC and anti-IgD-biotin were gifts from G. Klaus, and anti-CD25-biotin was a gift from R. Zamojska. Biotinylated antibodies were revealed with streptavidin-phycoerythrin (PE) (Biogenesis, Poole).

To study this defect further, we used mutant and control fetal liver cells to make radiation chimaeras with lethally irradiated B6 and sublethally irradiated RAG-1-deficient mice¹⁷. Analysis of bone marrow from the mutant B6 chimaeras showed that there were very few immature B cells (B220⁺IgM⁺IgD⁻) and virtually undetectable numbers of mature B cells (B220⁺IgM⁺IgD⁺) (Fig. 2c, and data not shown). To determine where the development of B cells was blocked, we examined the expression of the CD25 antigen in B220⁺IgM⁻ cells; this protein is upregulated during the transition from pro-B to pre-B cells^{18,19}. *syk*^{-/-} bone marrow contained normal numbers of B220⁺CD25⁺IgM⁻ pro-B cells, but substantially reduced numbers of B220⁺CD25⁺IgM⁻ pre-B cells (Fig. 2c, d). The pro- to pre-B cell transition requires the expression of a pre-B-cell receptor (pre-BCR), consisting of a functionally rearranged μ_H chain and the surrogate light chains λ_5 and V_{pre-B} (refs 20, 21). The pre-BCR is thought to deliver a signal that either permits the pro-B cell to mature into a pre-B cell or allows pre-B cells to survive and/or proliferate²². In view of the block at this transition in *syk*^{-/-} mice, we suggest that Syk may be important in signal transduction through the pre-BCR.

The spleen and lymph nodes of *syk*^{-/-} chimaeras contained no detectable conventional (IgM⁺IgD⁺) B cells (Fig. 2e). B-1



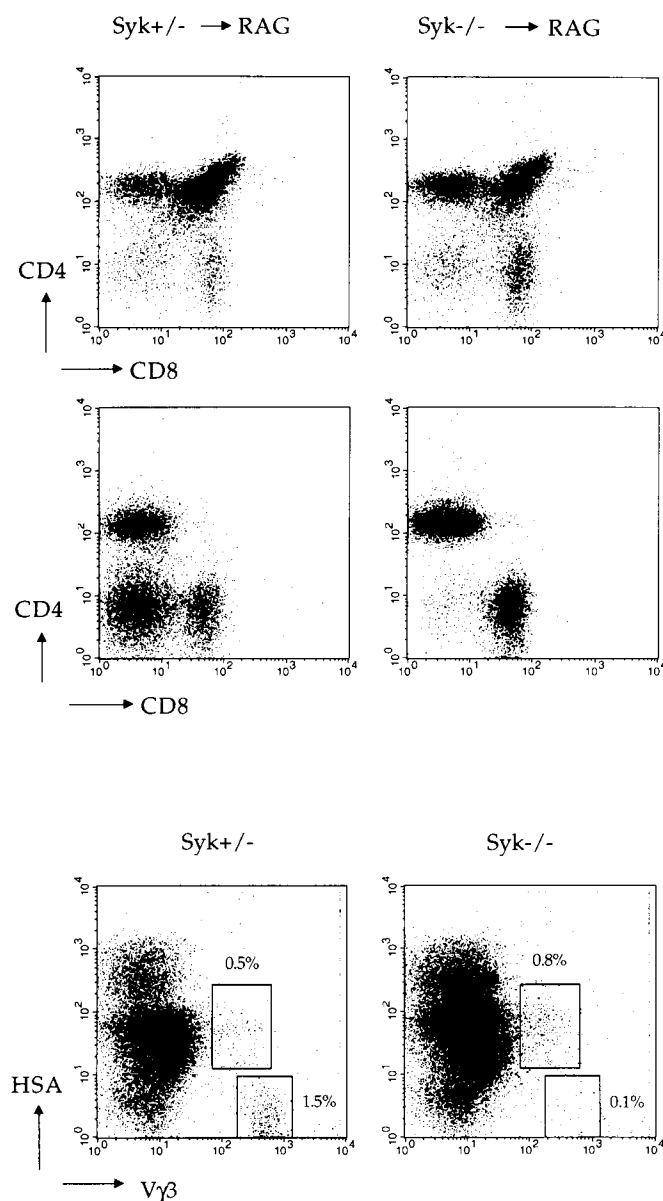


FIG. 3 T-cell development in *syk*^{-/-} mice. Flow cytometric analysis of staining for CD4 and CD8 of cells falling within the lymphocyte gate of thymus (top panels) and lymph nodes (middle panels) of RAG-1 radiation chimaeras reconstituted with *syk*^{+/-} and *syk*^{-/-} fetal liver. Bottom panels show staining of E16.5 thymuses from *syk*^{+/-} and *syk*^{-/-} fetuses for HSA and V γ 3. Immature (V γ 3^{lo}HSA⁺) and mature (V γ 3^{hi}HSA⁻) thymocytes are boxed. Percentages of cells falling within the boxes are shown as a fraction of all cells in the lymphocyte scatter gate. The *syk*^{-/-} thymuses contained significantly fewer mature V γ 3 thymocytes (*syk*^{+/-} $1.24 \pm 0.68\%$ ($n=5$); *syk*^{-/-} $0.23 \pm 0.12\%$ ($n=4$); $P < 0.003$). METHODS. Flow cytometric analysis was carried out as described for Fig. 2. Antibodies used were anti-CD4-PE (Boehringer-Mannheim), anti-CD8-APC, and anti-V γ 3-PE (536) (both from Pharmingen). Anti-HSA-biotin was a gift from R. Zamoyska. Biotinylated antibodies were revealed with streptavidin-QuantumRed (Sigma).

cells which carry IgM and low levels of IgD (IgM⁺IgD^{lo}) and predominate in the peritoneal cavity of control chimaeras, were also undetectable in the mutant chimaeras. Mice deficient in $\lambda 5$ also show a profound block at the pro- to pre-B cell transition and have reduced numbers of immature B cells²¹, but they accumulate mature conventional and peritoneal B cells^{21,23}. In contrast, despite the presence of immature B cells in the *syk*^{-/-}

radiation chimaeras, virtually no mature B cells were detected in the periphery up to 5 months following reconstitution; this indicates that Syk may be required for the transition of immature to mature B cells, their survival or subsequent expansion.

The lymphoid organs of *syk*^{-/-} chimaeras showed normal numbers of T cells in stains for CD4 and CD8 antigens and we found no differences in expression of CD3, $\alpha\beta$ T-cell receptor, CD69 and heat stable antigen (HSA), indicating that the development of *syk*^{-/-} $\alpha\beta$ T cells is normal (Fig. 3, and data not shown). *syk*^{-/-} T cells proliferated normally in response to anti-CD3, concanavalin A and IL-2, demonstrating that Syk is not essential for either T-cell receptor or IL-2 receptor-mediated signalling in mature T cells (not shown). Despite this, Syk may normally be involved in these signalling processes, but in *syk*^{-/-} T cells other kinases, such as ZAP-70, may substitute for it²⁴.

Analysis of E16.5 thymuses stained for V γ 3 and HSA showed that although *syk*^{-/-} thymuses generate immature V γ 3⁺ thymocytes (V γ 3^{lo}HSA⁺), they have very few mature V γ 3⁺ cells (V γ 3^{hi}HSA⁻) (Fig. 3). Analysis of thymuses at E15.5, E17.5 and E18.5 showed that mature V γ 3⁺ thymocytes are not detectable at any other time of development in either control or mutant mice (not shown). In contrast, the development of other $\gamma\delta$ -bearing T cells seems to be unaffected (not shown). Because V γ 3⁺ thymocytes exclusively become dendritic epidermal T cells²⁵, we analysed the epidermis of *syk*^{-/-} mice and found that these T cells were present, although only at 60% of the numbers in control mice (not shown). Thus the *syk* mutation probably does not alter the timing of V γ 3⁺ thymocyte development, but seems to affect their maturation, indicating a possible signalling role for Syk in these cells.

In conclusion, we have shown that Syk has a critical role in B-cell development. The related tyrosine kinase ZAP-70 is required for normal T-lymphocyte development and function^{8,9}. Together with our data, this suggests that this class of non-receptor tyrosine kinases is essential for antigen-receptor signal transduction. □

Received 1 June; accepted 8 September 1995.

1. Saouaf, S. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 9524–9528 (1994).
2. Takata, M. et al. *EMBO J.* **13**, 1341–1349 (1994).
3. Minami, Y. et al. *Immunity* **2**, 89–100 (1995).
4. Corey, S. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4683–4687 (1994).
5. Ravetch, J. V. *Cell* **78**, 553–560 (1994).
6. Couture, C., Baier, G., Altman, A. & Mustelin, T. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5301–5305 (1994).
7. Clark, E. A., Shattil, S. & Brugge, J. S. *Trends biochem. Sci.* **19**, 464–469 (1994).
8. Arpaia, E., Shahar, M., Dadi, H., Cohen, A. & Roifman, C. M. *Cell* **76**, 947–958 (1994).
9. Negishi, I. et al. *Nature* **376**, 435–438 (1995).
10. Shivdasani, R. A. et al. *Cell* **81**, 695–704 (1995).
11. Levéen, P. et al. *Genes Dev.* **8**, 1875–1887 (1994).
12. Soriano, P. *Genes Dev.* **8**, 1888–1896 (1994).
13. Herberstein, B. M. & Wallace, M. E. *J. med. Genet.* **1**, 10–23 (1964).
14. Hardy, R. R. & Hayakawa, K. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11550–11554 (1991).
15. Paige, C. J., Gislis, R. H., McKearn, J. P. & Iscove, N. N. *Eur. J. Immun.* **14**, 979–987 (1984).
16. Ogawa, M. et al. *EMBO J.* **7**, 1337–1343 (1988).
17. Spanopoulou, E. et al. *Genes Dev.* **8**, 1030–1042 (1994).
18. Rolink, A., Grawunder, U., Winkler, T. H., Karasuyama, H. & Melchers, F. *Int. Immun.* **6**, 1257–1264 (1994).
19. Chen, J., Ma, A., Young, F. & Alt, F. W. *Int. Immun.* **6**, 1265–1268 (1994).
20. Kitamura, D., Roes, J., Kühn, R. & Rajewsky, K. *Nature* **350**, 423–426 (1991).
21. Kitamura, D. et al. *Cell* **69**, 823–831 (1992).
22. Melchers, F. et al. *Curr. Opin. Immun.* **7**, 214–227 (1995).
23. Ehlich, A. et al. *Cell* **72**, 695–704 (1993).
24. Kong, G., Bu, J., Kurosaki, T., Shaw, A. S. & Chan, A. C. *Immunity* **2**, 485–492 (1995).
25. Havran, W. L. & Allison, J. P. *Nature* **344**, 68–70 (1990).
26. Hanks, S. K., Quinn, A. M. & Hunter, T. *Science* **241**, 42–52 (1988).
27. Tybulewicz, V. L., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. *Cell* **65**, 1153–1163 (1991).
28. von Freuden-Jeffry, U. et al. *J. exp. Med.* **181**, 1518–1526 (1995).
29. Guidos, C. J., Williams, C. J., Wu, G. E., Paige, C. J. & Danska, J. S. *J. exp. Med.* **181**, 1187–1195 (1995).
30. Ramsden, D. A., Paige, C. J. & Wu, G. E. *J. Immun.* **153**, 1150–1160 (1994).

ACKNOWLEDGEMENTS. We thank C. Atkins for flow cytometric analysis; G. Wu for advice on PCR analysis of immunoglobulin rearrangements; R. Zamoyska, A. Tarakhovskiy, L. Webb and G. Klaus for reviews of the manuscript; A. Pawson and A. Cheng for exchange of unpublished information; R. Gregory (Genzyme) for pS3C3; R. Zamoyska and G. Klaus for gifts of antibodies; D. Baltimore for RAG-1-deficient mice; D. Rennick (DNAX) for GM-CSF and G-CSF; and A. Fisher for ST2 cells. This work was supported in part by the Medical Research Council, and by grants from the Oliver Bird Fund (V.T.), the Leukemia Research Fund (V.T.) and the National Cancer Institute (R.L.G.).

Syk tyrosine kinase required for mouse viability and B-cell development

Alec M. Cheng*, Bruce Rowley†, William Pao‡, Adrian Hayday‡, Joseph B. Bolen‡ & Tony Pawson*§||

* Programme in Molecular Biology and Cancer, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Ave, Toronto, Ontario M5G 1X5, Canada

† Department of Oncology, Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, New Jersey 08453, USA

‡ Department of Biology, Yale University, KBT 616, PO Box 208103, New Haven, Connecticut 06520-8103, USA

§ Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada

THE Syk cytoplasmic protein-tyrosine kinase has two amino-terminal SH2 domains and a carboxy-terminal catalytic domain¹. Syk, and its close relative ZAP-70 (ref. 2), are apparently pivotal in coupling antigen- and Fc-receptors to downstream signalling events^{3,4}. Syk associates with activated Fc receptors⁵, the T cell receptor complex⁶ and the B-cell antigen-receptor complex (BCR) in immature and mature B lymphocytes⁷. On receptor activation, the tandem SH2 domains of Syk bind dual phosphotyrosine sites in the conserved ITAM motifs of receptor signalling chains, such as the immunoglobulin α and β -chains of the BCR, leading to Syk activation^{3,4,8}. Here we have investigated Syk function *in vivo* by generating a mouse strain with a targeted mutation in the *syk* gene. Homozygous *syk* mutants suffered severe haemorrhaging as embryos and died perinatally, indicating that Syk has a critical role in maintaining vascular integrity or in wound healing during embryogenesis. Analysis of *syk*^{-/-} lymphoid cells showed that the *syk* mutation impaired the differentiation of B-lineage cells, apparently by disrupting signalling from the pre-BCR complex and thereby preventing the clonal expansion, and further maturation, of pre-B cells.

An exon in the mouse *syk* gene was deleted in embryonic stem cells, and this mutation was transmitted to the germ line (Fig. 1a, b). The deleted exon encodes 41 residues that are located in subdomain VI of the Syk kinase domain, and are identical between the human and mouse proteins. No Syk protein was detected in lysates from *syk*^{-/-} fetal liver cells (Fig. 1c), suggesting that the engineered deletion interferes with Syk expression or stability and can therefore be considered a null mutation. No homozygous mutant mice were detected among 150 offspring at weaning age, indicating that Syk is essential for viability in the mouse. *syk*^{-/-} embryos at embryonic day 14 (E14) had severe systemic haemorrhage and oedema (Fig. 1d), which persisted until E18 and then gradually receded. Most mutants were born with red swollen footpads and a much reduced haemorrhage pattern, but generally died 1–5 days after birth. The systemic haemorrhage of *syk* mutants resembles that observed in mice deficient in platelet-derived growth factor B (PDGF-B) or its β -receptor^{9,10}. Syk is expressed in platelets¹¹, and is activated in human platelets by stimulation with thrombin or platelet-activating factor¹², or by fibrinogen binding to the α IIb β 3 integrin¹³. Whether the haemorrhage in *syk*^{-/-} embryos results from platelet dysfunction or has a more direct cause, such as an endothelial cell defect, is unknown. Syk-deficient fetal liver cells showed no defect in differentiation into macrophages, granulocytes or mast cells *in vitro* (data not shown). Analysis of blood from embryos and newborn mice revealed no overt defects in blood clotting, erythrocyte maturation or platelet numbers (data not shown).

To circumvent the perinatal lethality inflicted by the *syk* mutation, we assessed the ability of Syk-deficient haematopoietic stem cells from fetal liver to differentiate into T and B lymphocytes, after being transferred to immunodeficient *rag-2* mutant recipients¹⁴ that had been lethally irradiated. These *syk*^{-/-} cells differentiated *in vivo* into mature thymocytes carrying the CD4 or CD8 antigen and high numbers of T-cell receptors (CD4⁺TCR^{hi}, or CD8⁺TCR^{hi}; Fig. 2a), although the total number of thymocytes was usually lower than that found in *syk*^{+/-} reconstituted mice (by 28%, $n=5$, data not shown). In the spleen, both CD4⁺ and CD8⁺ T lymphocytes were present at a normal ratio for the CD4/CD8 population of 2:1 to 3:1. In this regard, Syk seems to be quite distinct from the T-cell specific ZAP-70 tyrosine kinase, which in humans is required for the development of CD8⁺ thymocytes from CD4⁺CD8⁺ precursors^{15,16}.

In contrast to the relatively normal reconstitution of T cells, *syk*^{-/-} haematopoietic stem cells failed to produce mature B220⁺IgM⁺ B cells in the spleen (Fig. 2a). A distinct lineage of peritoneal CD5⁺ B cells¹⁷ was also missing from mice reconstituted with Syk-deficient stem cells (Fig. 2a). In the bone marrow, most B-lineage cells were arrested at the CD43⁺ pro-B stage, and only a small population of the pre-B cells became B220^{hi}CD43^{lo}. In the most severe case, the CD43^{lo} population was completely absent, and no B220⁺IgM⁺ cells were detected (data not shown). Consistent with this surface marker phenotype, the number of cells in the small late pre-B population was drastically reduced (Fig. 2b), further indicating that B-cell development was arrested at the late pro-B stage^{18,19}. Although *rag-2*^{-/-} mice also show an arrest of B-cell development at the CD43⁺ stage¹⁴, Southern analysis indicated that the bone marrow population in the recipient mice was largely donor-derived (Fig. 2c). Furthermore, in contrast to the *rag-2* mutation, which arrests early B-cell development by blocking the recombination of immunoglobulin genes¹⁴, the *syk* mutation did not affect recombination between the genes for heavy (H)-chain diversity (D) and joining (J) segments and between the H-chain variable (V) segment and DJ segments (data not shown).

These data suggest that the maturation of Syk-deficient B-lineage cells is blocked at the transition between fractions C and D, defined according to the established model of B-cell development (Fig. 2d)^{18,19}. After the generation of a productively rearranged heavy (H)-chain gene, cells in fraction C express on their surface a pre-BCR, comprised of a membrane-bound μ -chain together with the V_{pre-B} and λ 5 surrogate light chains^{20,21}. It has been postulated that pre-B cells expressing the pre-BCR are clonally expanded before the activation of light-chain gene rearrangement^{18,22}, accounting both for the large proportion of pre-B cells bearing a productively rearranged H-chain gene, and for cell proliferation during the transition from the CD43⁺ to CD43⁻ stages^{18,19,23,24}.

To investigate further the possibility that pre-BCR signalling was disrupted in Syk-deficient pre-B cells, the ratio of productive to nonproductive H-chain genes in the Syk-deficient pre-B-cell population was assessed by population analysis using the polymerase chain reaction-restriction-fragment length polymorphism (PCR-RFLP) technique (Fig. 3). Greater than 65% of the rearranged H-chain genes in the sorted B220⁺IgM⁻ bone marrow B cells from wild-type mice, or from *rag-2*^{-/-} recipients of *syk*^{+/-} fetal liver cells, gave polymorphic H-chain gene fragments that were representative of in-frame products (Fig. 3a, lanes 5, 6). By comparison, V_H-DJ_H DNA fragments from *syk*^{-/-} pre-B cells displayed a pattern in which the in-frame bands were underrepresented, whereas the out-of-frame bands were overrepresented (Fig. 3b, compare lane 5 with lanes 6 and 7). The average frequency of in-frame products was about 36%, a value expected on the basis of random rearrangement. These data provide direct evidence that the developmental arrest caused by the *syk* mutation coincides with a lack of clonal expansion of cells expressing a functional μ -chain, and are consistent with a model in which the Syk tyrosine kinase mediates signalling downstream

|| To whom correspondence should be addressed at Samuel Lunenfeld Research Institute.

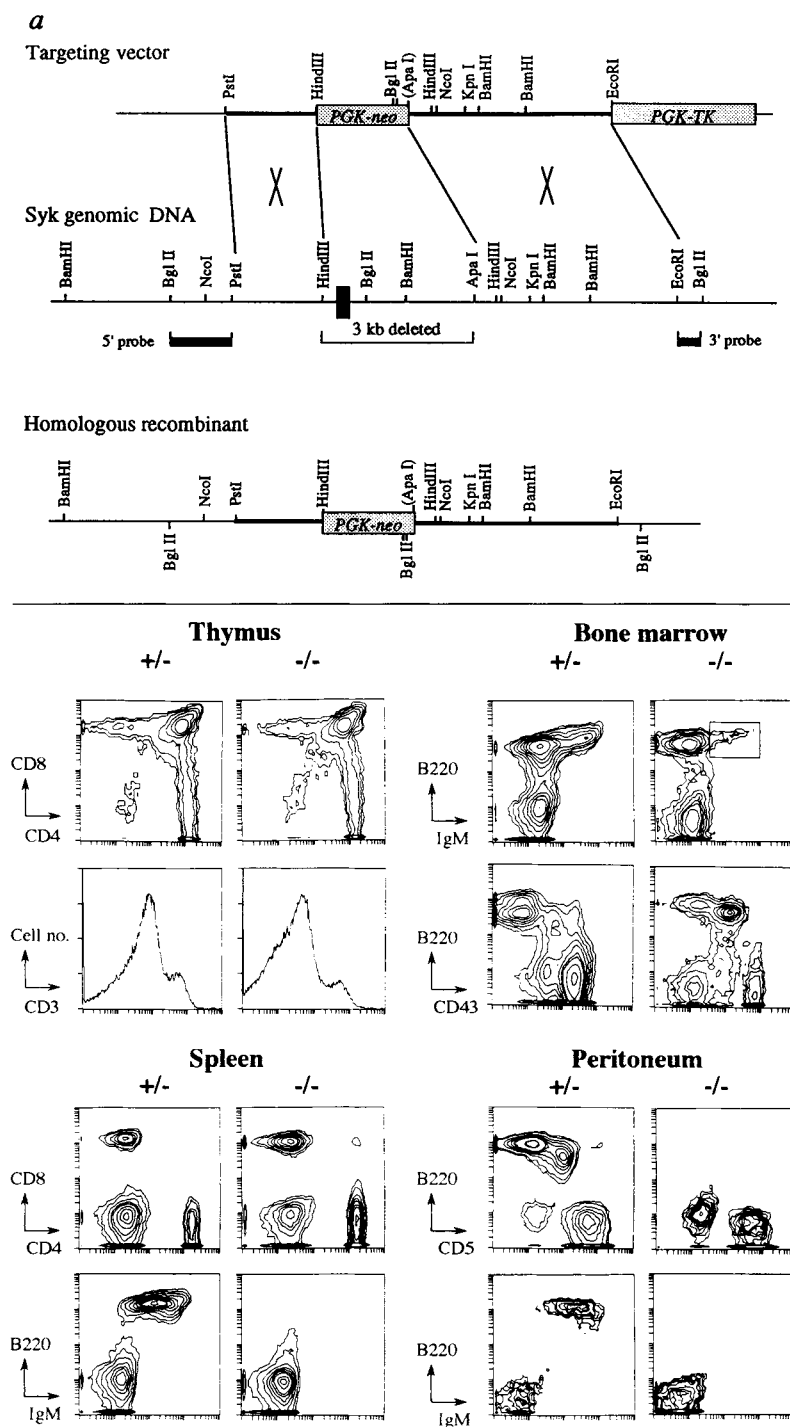
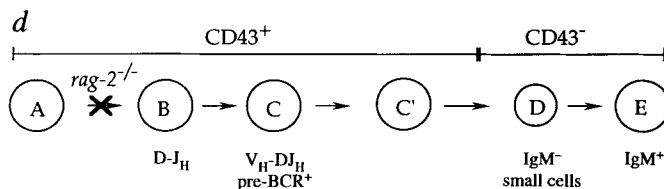
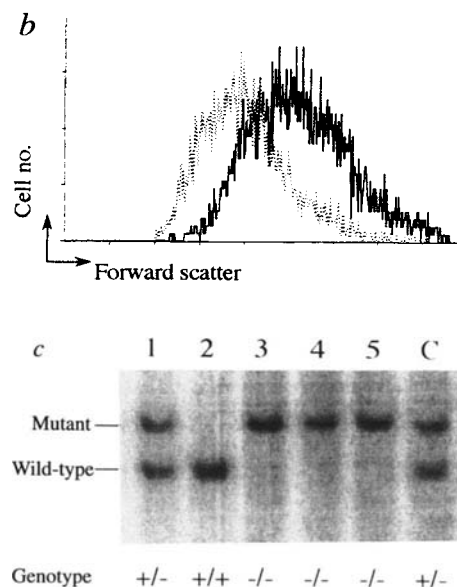


FIG. 2 Adoptive transfer of fetal liver haematopoietic stem cells in *rag-2*^{-/-} mice. **a**, Flow-cytometry analysis of single-cell suspensions from various lymphoid organs in *rag-2* mutant mice reconstituted with fetal liver cells of the indicated genotypes. Antibodies used for staining are indicated. The small fraction of bone marrow IgM⁺ B cells in Syk-deficient mice is framed. **b**, Cell size distribution of the bone marrow B220⁺ IgM⁺ population from *rag-2* mice transplanted with *syk*^{+/+} (solid dark line) or *syk*^{-/-} fetal liver cells (dotted line). Cell size correlates with the extent of forward scatter determined by flow-cytometry. **c**, Contribution of donor-derived stem cells in *rag-2*^{-/-} recipients. Genotypes of the donors are indicated (lanes 1–5). Control DNA from a heterozygous mutant mouse is included (lane C). **d**, Syk functions in bone marrow B-cell development. The major surface phenotypes, and the stages of D–J and V–DJ recombination of the H-chain locus, are shown. The points of developmental arrest induced by the *rag-2* mutation are indicated. The *syk* mutation blocks B-cell development in the transition from fractions C to D. **METHODS**. Fetal livers at 14–18 d.p.c. were disrupted by suction with a 3-ml syringe through a 21G needle to yield a single-cell suspension. Each liver was suspended in 1–2 ml of DME supplemented with 2% fetal

FIG. 1 Targeted disruption of the *syk* gene. **a**, Targeting strategy. The targeting vector containing the selection cassettes PGK-neo (neomycin) and PGK-TK (thymidine kinase) is shown. The deleted exon (black box) encodes 41 residues located in subdomain VI of the tyrosine kinase domain. In human Syk, this corresponds to residues 477–517. The left and right arms for homologous recombination are indicated by thickened lines. Predicted sizes of the *Bgl*II-digested genomic DNA fragments hybridized to the 5' and 3' probes were, respectively, 3.9 and 6.5 kb (endogenous) and 4.5 and 4.6 kb (recombinant). **b**, Southern blot analysis of neomycin- and gancyclovir-resistant embryonic stem (ES) cell clones. A Southern blot was hybridized with a 5' probe (lanes 1–5) and a 3' probe (lanes 6–10). The sizes of wild-type (3.9 kilobases kb, 6.5 kb) and recombinant (4.5 kb, 4.6 kb) bands are shown. **c**, Immunoblot analysis of fetal liver cells from embryos of the indicated genotypes with anti-Syk antiserum directed against the spacer region between the carboxyl-SH2 domain and the kinase domain (residues 260–370). The p72Syk (d.p.c. clones, genomic DNA was purified and digested with *Bgl*II directly on the 96-well plate. Digested DNAs were blotted onto protein is indicated. **d**, A *syk*^{-/-} embryo at 15 days postcoitum (d.p.c.) (left) is shown together with a normal *syk*^{+/+} embryo from the same litter.

METHODS. Two contiguous fragments of 3.5 and 7.0 kb containing an exon of *syk* encoding a conserved region of the tyrosine kinase domain were isolated from a 129/Sv mouse genomic library. A 3-kb region encompassing the conserved exon was targeted for deletion. R1 ES cells were electroporated with the targeting vector and



calf serum to give a final concentration of 4×10^6 to 1×10^7 cells ml⁻¹. *rag-2*^{-/-} mice at age 4–8 weeks were treated with a lethal dose of γ -irradiation (800 rad). Approximately 2.5×10^6 cells were transferred to each irradiated mouse by intravenous injection in the tail. Antibodies for flow cytometry were from Pharmingen. Cells were stained with specific antibodies and analysed with a Coulter-Elite cell sorter. To determine the contribution of bone marrow cells by fetal liver-derived stem cells, genomic DNA was isolated from total bone marrow cells at 4–8 weeks post-reconstitution, digested with *Bgl*II and analysed for the *syk* mutant allele by Southern blotting with a 5' probe as described for Fig. 1.

► selected on G418 and gancyclovir. Individual double-resistant clones were isolated and propagated on 96-well plates. To screen for homologously targeted clones, genomic DNA was purified and digested with *Bgl*II directly on the 96-well plate. Digested DNAs were blotted onto nitrocellulose filters and hybridized with the 5' or 3' probe. Two independent targeted lines were isolated and used to generate chimaeric mice as described²⁸. For detection of Syk protein, immunoprecipitation with Syk antiserum was carried out on lysates of fetal liver cell suspensions, isolated from 14 d.p.c. embryos. Proteins were separated in an 8% denaturing SDS-polyacrylamide gel, transferred to nitrocellulose, blotted with anti-Syk antiserum and detected by chemiluminescence.

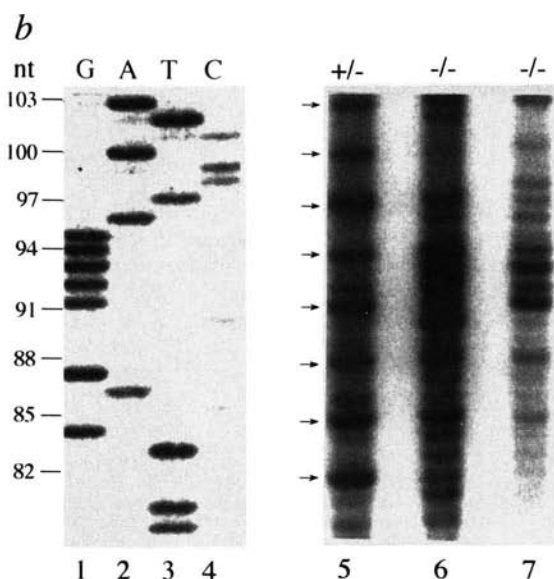
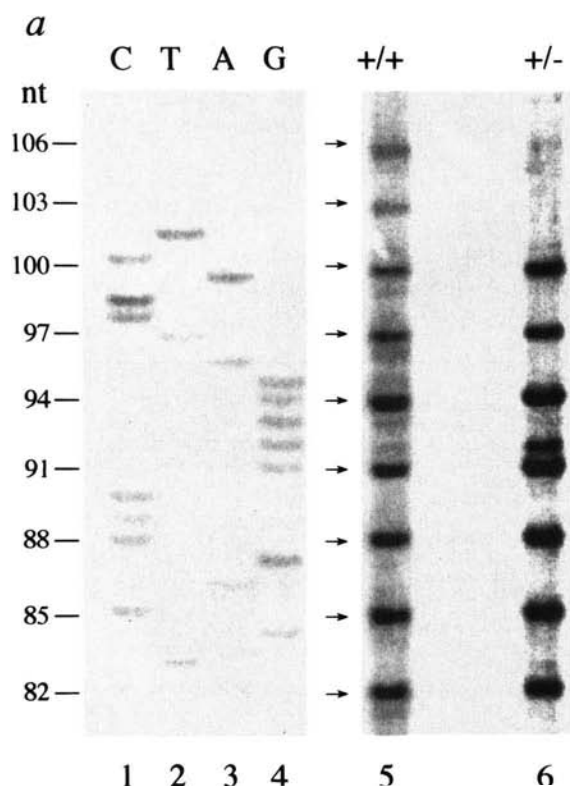
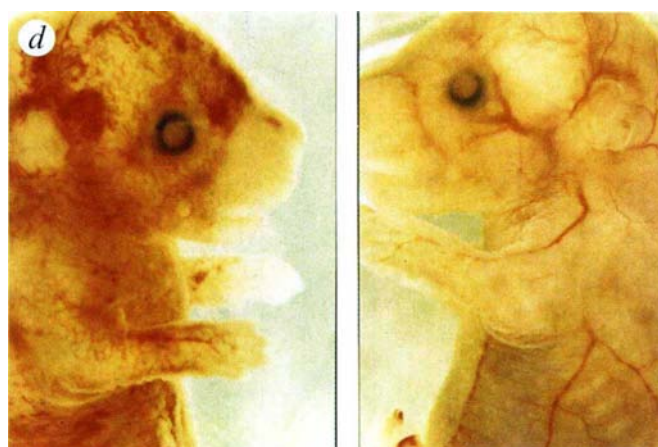
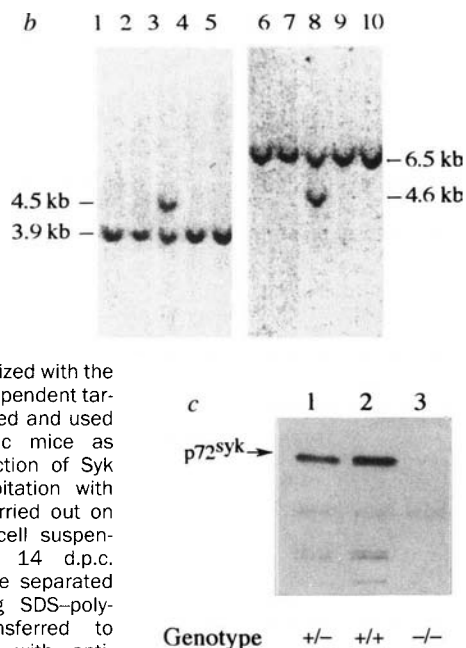
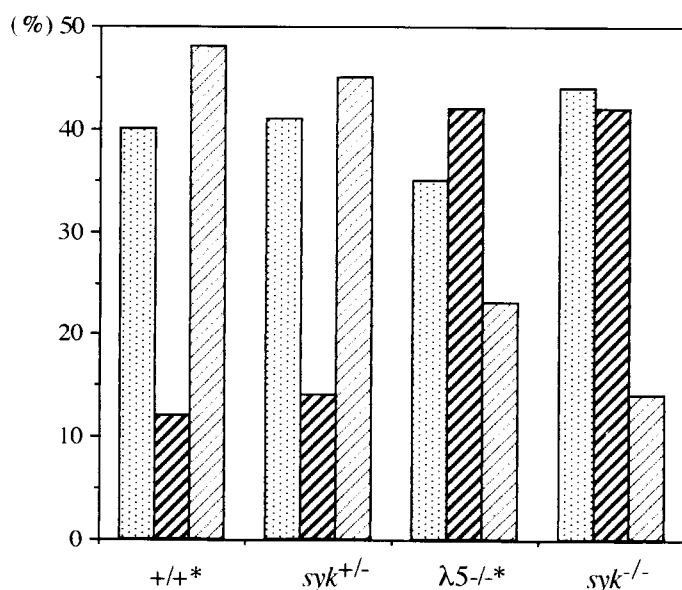


FIG. 3 Population analysis of productively and nonproductively rearranged immunoglobulin H-chain genes. *a*, Comparison of RFLP patterns in wild-type (lane 5) and *syk*^{-/-} (lane 6) bone marrow pre-B (B220⁺ IgM⁻) cell DNA. The sizes of DNA fragments in nucleotides (nt) are determined from the sequencing ladder as shown (lanes 1–4). Positions of in-frame bands are indicated by arrows. *b*, RFLP pattern in *syk*^{+/-} (lane 5) and *syk*^{-/-} pre-B cells (lanes 6, 7).

METHODS. PCR-RFLP analysis²⁹ was carried out on genomic DNA from sorted B220⁺ IgM⁻ bone marrow cells with the following modifications. *Pfu* polymerase (Stratagene) was used for DNA amplification, because it lacks significant terminal transferase activity. Amplified DNA fragments corresponding to recombination of *V_H7183/J_H4* were isolated. Digestion of PCR products with *Hae*III (see below) generated polymorphic frag-

ments spanning the region of diversity that have a 5'-phosphate at the end generated by the restriction enzyme, and a 5'-hydroxyl group at the end created by the *V_H7183* PCR primer. The 5'-OH termini were labelled by T4 polynucleotide kinase (NEB). The following primers were used for DNA amplification. *V_H7183* 5'-CGCGCGGCCGCTGGAGTCTTGGGGGAG-GCTA-3' and *J_H4* 5'-GCGCTCGAGGAGACGGTGACTGAGGTT-3'. The *V_H7183* oligonucleotide hybridizes to all 19 identified members of the *V_H7183* family^{30,31}, 11 of which contain (at amino acids 87–88) an internal *Hae*III site. An in-frame fusion of *V_H* to *J_H4* in these 11 genes would yield fragments of 79 base pairs (bp). Productive polymorphic fragments would occur at intervals of 3 nucleotides thereafter (for example, 82, 85, 88). Sizes of undigested and nonpolymorphic fragments were noticeably outside the range of the pieces analysed here. Bands were visualized next to a pBSSK ladder (sequenced with ³⁵S-dATP and T7 primer) following exposure to phosphorimaging plates (Molecular Dynamics). The frequency of the productive bands in each sample was quantified using the phosphorimager. It was noted previously that nonproductively rearranged products of a member of this family, *V_H7183.1*, were present at an unusually high rate²². However, this member does not contain the *Hae*III site and its rearrangements are not detected in this PCR-RFLP assay.



of the pre-BCR during B-cell maturation²⁵ (Fig. 2d). The *syk* mutation does not completely arrest bone marrow B-cell development (as shown by the CD43^{lo} population; Fig. 2a), a phenomenon also observed in the $\lambda 5$ -deficient mice¹⁹ and which may be due to the presence of a pre-BCR-independent pathway. As no mature *syk*^{-/-} B cells accumulate in the periphery, this observation argues that Syk must also govern the complete maturation of B cells, in addition to its role at the C-D transition. This might involve the control of cell survival, or migration of IgM⁺ B cells to the peripheral lymphoid organs.

During early B-cell development, pro-B cells in which D-J_H rearrangement occurs in reading frame II (D-rf II) express a membrane-bound protein, D μ . Normally, these cells are significantly underrepresented (approximately 14%, in contrast to 40% in rf I or rf III), suggesting that there is a selection against rf II

FIG. 4 Syk regulates signalling downstream of D μ . Reading frame (rf) distributions in *syk*^{+/+} and *syk*^{-/-} pre-B population are plotted in comparison with those previously reported²⁷ for wild-type and $\lambda 5$ deficient mice (*). $\square$ rf I; $\square$ rf II; $\square$ rf III.

METHODS. Genomic DNAs prepared from bone marrow B220⁺ IgM⁺ B cells were used as templates for PCR amplification. The sequences of primers specific for D and J_H4 segments and the PCR amplification profile have been described previously²⁶. Amplified fragments were gel-purified, end-polished with T4 DNA polymerase and subcloned into pBluescript (SK) at the SmaI site in the polylinker region. To avoid sampling bias, D-J_H4-positive recombinants were isolated without the aid of an X-gal marker, because D-J_H4 inserts produced both blue and white colonies. The DNA sequences of the inserts were determined with a primer specific for the 3' end to the J_H4 gene segment. A 1:15 dilution of dGTP mix was used to ensure detection of sequences immediately downstream of the 3' end of the primer (Sequenase Kit; USB). Rf distributions from the *syk*^{+/+} and *syk*^{-/-} B-cells were determined from 29 and 55 subclones, respectively.

usage in pro-B cells^{26,27}. This might occur through the induction of cell death. Alternatively, D μ signalling may activate the allelic exclusion machinery that inhibits V_H DJ_H recombination, thereby preventing μ -chain expression, and arresting further maturation. Strikingly, the *syk* mutation blocked D μ signalling (Fig. 4), as judged by an elevated frequency of D μ ⁺ cells (an increase in rf II frequency to >40%, compared with 14% in *syk*^{+/+} cells). This phenomenon also occurs in the $\lambda 5$ -deficient B cells²⁷.

The Syk tyrosine kinase therefore has multiple functions in the mouse. It is required in embryogenesis to prevent haemorrhaging, and at several discrete stages in the differentiation of B-lineage cells. These observations underscore the importance of SH2-mediated interactions and tyrosine kinase function in mammalian cellular development. $\square$

Received 6 June; accepted 8 September 1995.

1. Taniguchi, T. et al. *J. Biol. Chem.* **266**, 15790–15796 (1991).
2. Chan, A. C., Iwashima, M., Turk, C. W. & Weiss, A. *Cell* **71**, 649–662 (1992).
3. DeFranco, A. L. *Curr. Opin. Cell Biol.* **7**, 163–175 (1995).
4. Takata, M. et al. *EMBO J.* **13**, 1341–1349 (1994).
5. Durden, D. L. & Liu, Y. B. *Blood* **84**, 2102–2108 (1994).
6. Chan, A. C. et al. *J. Immunol.* **152**, 4758–4766 (1994).
7. Hutchcroft, J. E., Harrison, M. L. & Geahlen, R. L. *J. Biol. Chem.* **267**, 8613–8619 (1992).
8. Rowley, R. B., Burkhardt, A. L., Chao, H.-G., Matsueda, G. R. & Bolen, J. B. *J. Biol. Chem.* **270**, 11590–11594 (1995).
9. Leveen, P. et al. *Genes Dev.* **8**, 1875–1887 (1994).
10. Soriano, P. *Genes Dev.* **8**, 1888–1896 (1994).
11. Taniguchi, T. et al. *J. Biol. Chem.* **268**, 2277–2279 (1993).
12. Rezaul, K., Yanagi, S., Sada, K., Taniguchi, S. & Yamamura, H. *Fibrosis Hemostas.* **72**, 937–941 (1995).
13. Clark, E. A., Shattil, S. J., Ginsberg, M. H., Bolen, J. & Brugge, J. S. *J. Biol. Chem.* **269**, 28859–28864 (1994).
14. Shinkai, Y. et al. *Cell* **68**, 855–867 (1992).
15. Arpaia, E., Shahar, M., Dadi, H., Cohen, A. & Roifman, C. M. *Cell* **76**, 947–958 (1994).
16. Chan, A. C. et al. *Science* **264**, 1599–1601 (1994).
17. Hayakawa, K., Hardy, R. R., Herzenberg, L. A. & Herzenberg, L. A. *J. exp. Med.* **161**, 1554–1568 (1985).
18. Hardy, R. R., Carmack, C. E., Shinton, S. A., Kemp, J. D. & Hayakawa, K. *J. exp. Med.* **173**, 1213–1225 (1991).

19. Kitamura, D. et al. *Cell* **69**, 823–831 (1992).
20. Misener, V., Downey, G. P. & Jongstra, J. *Int. Immun.* **3**, 1129–1136 (1991).
21. Karasuyama, H., Kudo, A. & Melchers, F. *J. exp. Med.* **172**, 969–979 (1992).
22. Decker, D. J., Boyle, N. E., Koziol, J. & Klinman, N. R. *J. Immunol.* **146**, 350–359 (1991).
23. Spanopoulou, E. et al. *Genes Dev.* **8**, 1030–1042 (1994).
24. Young, F. et al. *Genes Dev.* **8**, 1043–1057 (1994).
25. Papavasiliou, F., Misulovin, Z., Suh, H. & Nussenzweig, M. C. *Science* **268**, 408–411 (1995).
26. Gu, H., Kitamura, D. & Rajewsky, K. *Cell* **65**, 47–54 (1991).
27. Haasner, D., Rolink, A. & Melchers, F. *Int. Immun.* **6**, 21–30 (1994).
28. Joyner, A. L. in *Gene targeting. A practical approach* (IRL, Oxford, 1993).
29. Mallick, C. A., Dudley, E. C., Viney, J. L., Owen, M. J. & Hayday, A. C. *Cell* **73**, 513–519 (1993).
30. Carlsson, L., Overmo, C. & Holmberg, D. *Eur. J. Immun.* **22**, 71–78 (1992).
31. Chukwuocha, R. U., Hartman, A. B. & Feeney, A. J. *Immunogenet.* **40**, 76–78 (1994).

ACKNOWLEDGEMENTS. We thank F. Alt for the gift of Rag-2-deficient mice; C. Roman and D. Baltimore for discussion and PCR primers used in analysis of B220, $\lambda 5$, V_{pre-B}, immunoglobulin- α and - β expression and κ -light-chain gene rearrangement; C. Paige and S. Anderson for helpful discussions; R. Paulson and E. Smith for critical reading of the manuscript; and S. Gardner for generation of mouse chimeras. This work was supported by a grant from Bristol-Myers Squibb and by an International Research Scholar Award from the Howard Hughes Medical Institute to T.P.; A.C. is a postdoctoral fellow of the Leukemia Research Foundation; W.P. has been supported by Developmental Biology Training (NIH) and MSTP (NIH) grants; A.H. is supported by a NIH grant; T.P. is a Terry Fox Cancer Research Scientist of the National Cancer Institute of Canada.

Phosphorylation of Raf by ceramide-activated protein kinase

Bei Yao, Yuhua Zhang, Sylvie Dellkat,
Shalini Mathias, Subham Basu
& Richard Kolesnick*

Laboratory of Signal Transduction,
Memorial Sloan-Kettering Cancer Center, New York,
New York 10021, USA

THE sphingomyelin pathway, initiated by hydrolysis of sphingomyelin to ceramide and stimulation of a Ser/Thr ceramide-activated protein (CAP) kinase, mediates tumour necrosis factor- α (TNF- α) and interleukin-1 β action¹⁻⁴. CAP kinase is membrane-bound and proline-directed, recognizing the minimal substrate motif Thr-Leu-Pro (ref. 5). TNF may use the sphingomyelin pathway to signal Raf1 to activate the MAP kinase cascade⁶⁻⁸. Evidence shows that cytoplasmic Raf1 binds to GTP-ras upon cellular stimulation, is recruited to the plasma membrane, and activated^{9,11}. How membrane-bound Raf1 is activated is uncertain, but regulation of its kinase activity may involve its phosphorylation¹²⁻¹⁹. Specific Raf kinases, however, have not hitherto been identified. Here we report that CAP kinase phosphorylates Raf1 on Thr 269, increasing its activity towards MEK (MAP kinase or ERK kinase). Moreover, in intact HL-60 cells, CAP kinase complexes with Raf1 and, in response to TNF and ceramide analogues, phosphorylates and activates Raf1, implicating CAP kinase as a link between the TNF receptor and Raf1.

CAP kinase from HL-60 cells was resolved by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and renatured as a 97K protein (see legend to Fig. 1)². Under *in vitro* conditions, there was minimal autophosphorylation of recombinant Raf1 in the absence of CAP kinase (Fig. 1a). Further, CAP kinase isolated from unstimulated cells showed little activity towards Raf1. Raf1 phosphorylation, however, was enhanced by CAP kinase derived from ceramide- or sphingomyelinase-treated cells (Fig. 1a), as well as from TNF-stimulated cells (Fig. 1b). Purified CAP kinase from bovine brain gave similar results (data not shown). Thus Raf1 serves as a CAP kinase substrate *in vitro*, and CAP kinase activity towards Raf1 is increased by TNF stimulation.

Raf1, pretreated with CAP kinase, was fourfold more active in phosphorylating MEK1 than untreated Raf1, indicating that phosphorylation of Raf1 enhanced its kinase activity (Fig. 1c). CAP kinase did not phosphorylate MEK1 directly (Fig. 1d). Raf1 is therefore activated by phosphorylation by CAP kinase. In this regard, CAP kinase may require a Raf1 substrate possessing some baseline activity as one recombinant preparation lacking activity was not activated (data not shown).

We reconstituted the entire MAP kinase cascade²⁰ from CAP kinase to MAP kinase *in vitro* with purified reagents (Fig. 1e). Under our assay conditions, Raf1 and MEK1 together increased MAP kinase phosphorylation and enhanced MAP kinase activity fivefold. Addition of CAP kinase to these incubations further increased MAP kinase phosphorylation and activity 30-fold over control. Because CAP kinase failed to activate MEK1 or MAP kinase directly (Fig. 1d, e), these studies show that activation of MAP kinase by CAP kinase required Raf1.

In the next series of experiments, endogenous Raf1 immunoprecipitated from resting HL-60 cells was phosphorylated by CAP kinase for use in the MAP kinase reconstitution assay (Fig. 1f). Raf1 from resting HL-60 cells had little effect, but Raf1 prephosphorylated by CAP kinase, like the recombinant sub-

strate, enhanced signalling through to MAP kinase 20-fold. Moreover, dephosphorylation of CAP kinase-treated Raf1 with potato acid phosphatase abolished MEK1 phosphorylation and signalling through to MAP kinase (data not shown).

Tryptic digestion of CAP kinase-treated Raf1 generated one major radioactive peak phosphorylated exclusively on threonine residues (Fig. 2a, b). Edman degradation (Fig. 2b) revealed that CAP kinase phosphorylates Raf1 on threonine 268 and 269, with a slight preference for Thr 269 as the phosphoacceptor site. This latter site is contained within a -TLP- (single-letter amino-acid code) motif known as the preferred recognition site for CAP kinase⁵. Further, we have found that CAP kinase can phosphorylate peptides derived from the amino-acid sequence surrounding Thr 268 and 269 of Raf1 (amino acids 254-278). Bovine brain CAP kinase phosphorylated a peptide containing the wild-type Raf1 sequence TTLP (Fig. 2c), with phosphorylation occurring exclusively on threonine residues (data not shown). In contrast, CAP kinase failed to phosphorylate a peptide in which Thr 268 and 269 were substituted with alanines, generating the motif AALP. Additional studies using peptides containing the motifs ATLP and TALP showed that replacement of Thr 269 with alanine also abolished phosphorylation on Thr 268, whereas replacement of Thr 268 did not affect phosphorylation of Thr 269. A Raf1 mutant in which threonine 268 and 269 were substituted with valine residues retained some basal activity towards MEK (32% of wild-type Raf1), but was not a substrate for CAP kinase (data not shown). Neither did it support CAP kinase-induced activation of the MAP kinase cascade *in vitro* (Fig. 2d).

In intact HL-60 cells, Raf1 phosphorylation is induced seconds after stimulation with TNF (Fig. 3), consistent with the notion that Raf1 might be a component of the TNF signalling pathway^{7,8}. The effect, which persisted for at least 20 min, could be similarly induced by ceramide (25 μ M) and sphingomyelinase (10 mU ml⁻¹; data not shown). Treatment of the cells with TNF also enhanced the activity of Raf1 towards MEK1 10-20-fold (Fig. 3b). Raf1 derived from ceramide- or sphingomyelinase-treated cells possessed similarly enhanced activity towards MEK1 (Fig. 3c). Dephosphorylation of immunoprecipitated Raf1 with potato acid phosphatase abolished this enhanced activity (data not shown). Raf1 from treated cells not only phosphorylated MEK1 but was also 25-fold more active in our reconstituted MAP kinase signalling assay (Fig. 3d). These studies indicate that Raf1 is a component of the sphingomyelin pathway mediating TNF action.

We also compared the effect of TNF to other agents that activate Raf1. The phorbol ester 12-*O*-tetradecanoylphorbol 13-acetate (TPA; 100 ng ml⁻¹) did not activate CAP kinase yet increased Raf1 activity to fourfold that of the control after 10 min. Insulin (100 nM) similarly increased Raf1 activity in cells preincubated in serum-free medium as described²¹. Granulocyte/macrophage colony-stimulating factor (500 pM) only enhanced Raf1 activity 2.5-fold, whereas platelet-derived growth factor (PDGF) (5 nM) treatment of HL-60 cells, after induction of PDGF receptors with TPA²², resulted in a fourfold increase in Raf1 activation. Hence, the effect of TNF on Raf1 activity in HL-60 cells appears larger than that of these other agonists.

Raf1 may form a multiprotein complex^{10,23} including heat-shock proteins, MEK, 14-3-3 proteins^{24,26} and, in some cells, the EGF receptor²⁷ and Bcl2 (ref. 28). Raf1, immunoprecipitated from TNF-treated HL-60 cells, was therefore subjected to an immune-complex kinase assay. Figure 4a shows that a spectrum of proteins immunoprecipitated with Raf1 becomes phosphorylated *in vitro* under these conditions. These other proteins were not directly recognized by the primary anti-Raf1 antibody in western blots (Fig. 4b), and were not immunoprecipitated by a nonspecific antibody (data not shown). TNF treatment of cells resulted in enhanced phosphorylation of numerous proteins within this complex. A barely detectable band of 97K in control incubations was readily observed in Raf1 immunoprecipitates

* To whom correspondence should be addressed.

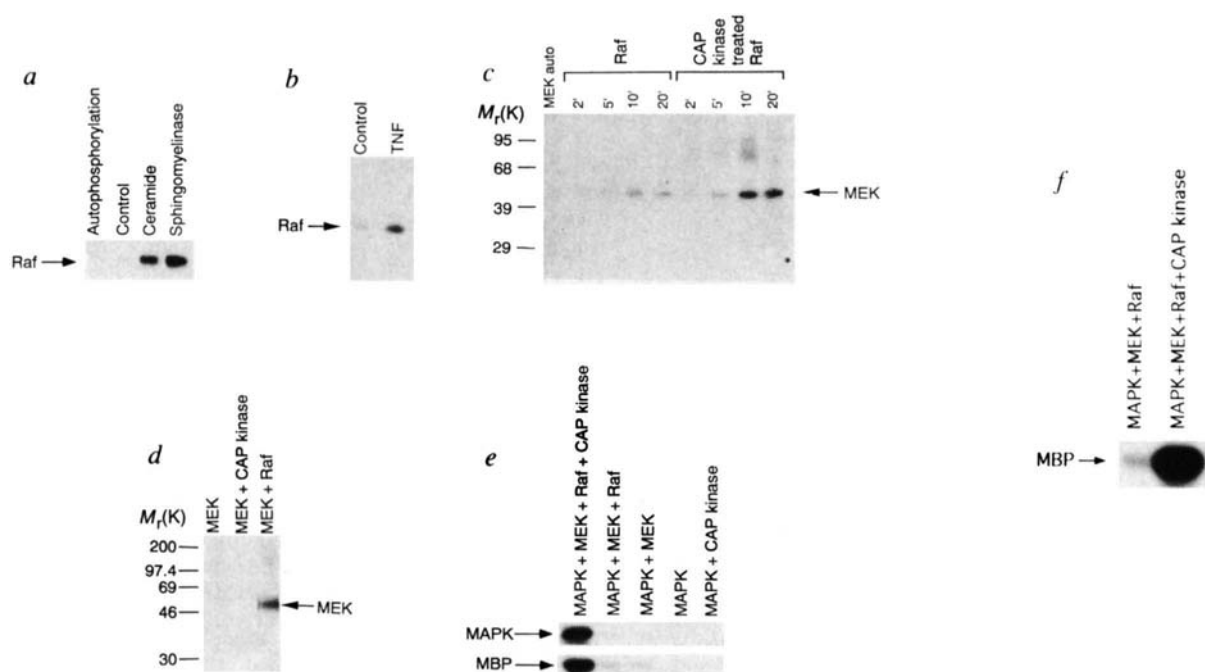


FIG. 1 CAP kinase phosphorylates recombinant human Raf1 *in vitro* activating the MAP kinase cascade. **a**, CAP kinase was renatured from untreated (Control), and C8-ceramide- and *Staphylococcus aureus* sphingomyelinase-treated HL-60 cells and used to phosphorylate recombinant human Raf1. **b**, CAP kinase, renatured from untreated (Control) and TNF-stimulated HL-60 cells, was used to phosphorylate recombinant Raf1. **c**, Raf1, phosphorylated by CAP kinase, has enhanced kinase activity towards MEK1. **d**, CAP kinase does not phosphorylate MEK1. **e**, Reconstitution of the MAP kinase cascade *in vitro* using recombinant Raf1. **f**, CAP kinase activates Raf1 immunoprecipitated from HL-60 cells.

METHODS. **a**, HL-60 cells (1×10^6 ml $^{-1}$), incubated in serum-free RPMI for 2 h, were stimulated with C8-ceramide (25 μ M) or *S. aureus* sphingomyelinase (Boehringer; 10 mU ml $^{-1}$). After 20 min, cells were homogenized in buffer (25 mM HEPES, pH 7.4, 5 mM EGTA, 50 mM NaF containing 10 μ g ml $^{-1}$ soy bean trypsin inhibitor and leupeptin), microsomal membranes prepared, and CAP kinase (3×10^7 cell equivalents per lane) resolved by 7.5% SDS-PAGE and renatured as described². Briefly, the acrylamide gel harbouring CAP kinase was fixed in buffer A (50 mM Tris, pH 7.4, 5 mM β -mercaptoethanol) containing 20% isopropanol, denatured in buffer A containing 6 M guanidine HCl and renatured in buffer A containing 0.04% Tween-20 at 4 °C. To phosphorylate recombinant human Raf1 immunoprecipitated from lysates of Sf9 cells coexpressing Raf1, p21^{ras}, and activated pp60^{src} as described¹⁸, blank or CAP kinase-containing gel slices were mixed with Raf-bound beads in 40 μ l Raf reaction buffer (30 mM HEPES, pH 7.4, 5 mM MgCl₂, 10 mM MnCl₂, 1 mM dithiothreitol) containing 5 μ M ATP and 20 μ Ci [γ -³²P]ATP. After 30 min, phosphorylated Raf1 was resolved by 7.5% SDS-PAGE. Raf1 phosphorylation was linear for 30 min. A FLAG-tagged Raf1 was an equally effective substrate. These data represent one of three similar experiments. **b**, Experiments were performed as in **a** except cells received TNF (1 nM, 20 min). These data represent one of four similar experiments. CAP kinase renatured from cells stimulated for 5 min with TNF yielded identical results. **c**, Recombinant Raf1 was phosphorylated for 30 min with CAP kinase renatured from TNF-stimulated HL-60 cells as in **a**. Controls (Raf) received blank gel pieces. Raf1 activity was measured by phosphorylation of recombinant human MEK1 (0.1 μ g per reaction) in 50 μ l kinase reaction buffer (40 mM Tris, pH 7.5, 30 mM NaCl, 10 mM MgCl₂) containing 100 μ M ATP and 50 μ Ci [γ -³²P]ATP. Phosphorylated MEK1 was resolved by 10% SDS-PAGE. MEK1 autophosphorylation (MEK auto) was performed for 20 min. This figure represents one of five similar experiments. **d**, MEK1 autophosphorylation and phosphorylation by Raf1 or CAP kinase from TNF-treated cells were performed for 1 h as in **a** and **c**. **e**, A 'heavy' microsomal membrane fraction enriched 10-fold in plasma membrane was generated according to ref. 30 from a postnuclear supernate of 800 g bovine brain prepared as in **a**. This contained virtually all the cellular CAP kinase and was subfractionated over a discontinuous sucrose density gradient. The CAP kinase-enriched fraction was extracted with

1 M KCl, precipitated with ammonium sulphate, eluted from a FPLC hydroxyapatite column with phosphate buffer, and CAP kinase activity was resolved completely using a Prep Cell (BIO-RAD). Purified renatured bovine brain CAP kinase or blank gel pieces were incubated with or without recombinant FLAG/Raf1 and MEK1 in Raf reaction buffer containing 5 μ M ATP. After 30 min, CAP kinase was removed by centrifugation. To measure MAP kinase phosphorylation, agarose-conjugated human GST-MAP kinase (6.25 μ g per reaction, UBI) was added to the supernatant in kinase reaction buffer containing 50 μ M ATP and 50 μ Ci [γ -³²P]ATP and after 20 min was separated by centrifugation. To measure MAP kinase activity, MAP kinase was phosphorylated with unlabelled ATP and then incubated in kinase reaction buffer with 50 μ g MBP, 50 μ M ATP and 50 μ Ci [γ -³²P]ATP for 20 min. ³²P-labelled MAP kinase and MBP were resolved by 12% SDS-PAGE. CAP kinase from HL-60 cells yielded similar results. **f**, Reconstitution of the signalling pathway was as described in **e** except that endogenous Raf1, immunoprecipitated from 40×10^6 HL-60 cells as in **a**, was used as the substrate.

FIG. 2 Mapping of the site of Raf1 phosphorylation by CAP kinase. **a**, Reverse-phase HPLC analysis of ³²P-labelled phosphopeptides from a tryptic digest of Raf1 that had been phosphorylated by CAP kinase. ³²P radioactivity in each fraction is shown as counts per minute (c.p.m.). **b**, Edman degradation (left) and phosphoamino-acid analysis (PAA, right) of the tryptic phosphopeptide isolated in HPLC fraction 29 (shown in **a**). ³²P radioactivity released during each cycle of Edman degradation is shown. The amino-acid sequence of the peptide containing threonine 268 and threonine 269 (underlined) is shown. S, Phosphoserine; T, phosphothreonine; Y, phosphotyrosine. **c**, Phosphorylation by CAP kinase of Raf1 peptides derived from the site surrounding Thr 268 and Thr 269. **d**, Reconstitution of the MAP kinase cascade by wild-type and mutant Raf. 1

METHODS. **a**, FLAG/Raf1 was phosphorylated by CAP kinase as in Fig. 1, subjected to tryptic digestion, adjusted to pH 2 with 20% trifluoroacetic acid (TFA), and tryptic peptides were separated by a C₁₈ reverse-phase HPLC column (Waters 3.9 $\times$ 300 mm) using an increasing gradient of acetonitrile in 0.05% TFA as described¹⁸. Fractions were Cerenkov-counted for ³²P content. **b**, Semi-automated amino-terminal sequence analysis using a Beckman 890C spinning cup sequencer and phosphoamino-acid analysis of the peptide in HPLC fraction 29 were as previously described¹⁸. **c**, Purified bovine brain CAP kinase was renatured as described in Fig. 1 and used to phosphorylate Raf1 peptides. Peptides derived from the amino-acid sequence surrounding Thr 269 of Raf1 (amino acids 254–278) containing the wild-type sequence TTLP and the altered sequence AALP were synthesized using an Applied Biosystems model 431A synthesizer and used in substrates in the CAP kinase assay. These peptides were made slightly longer than the natural tryptic peptide from intact Raf1 corresponding to amino acids 257–

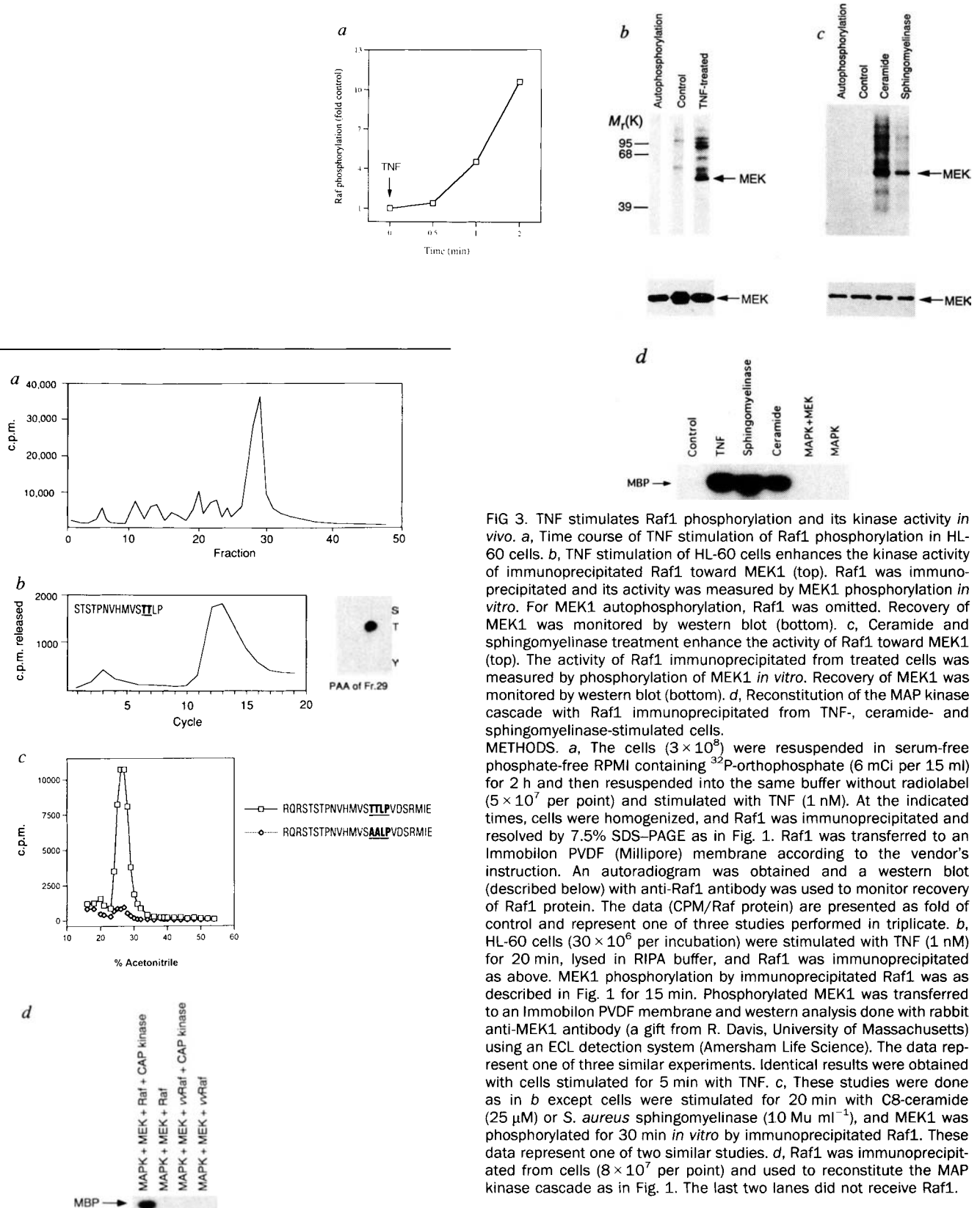


FIG 3. TNF stimulates Raf1 phosphorylation and its kinase activity *in vivo*. **a**, Time course of TNF stimulation of Raf1 phosphorylation in HL-60 cells. **b**, TNF stimulation of HL-60 cells enhances the kinase activity of immunoprecipitated Raf1 toward MEK1 (top). Raf1 was immunoprecipitated and its activity was measured by MEK1 phosphorylation *in vitro*. For MEK1 autophosphorylation, Raf1 was omitted. Recovery of MEK1 was monitored by western blot (bottom). **c**, Ceramide and sphingomyelinase treatment enhance the activity of Raf1 toward MEK1 (top). The activity of Raf1 immunoprecipitated from treated cells was measured by phosphorylation of MEK1 *in vitro*. Recovery of MEK1 was monitored by western blot (bottom). **d**, Reconstitution of the MAP kinase cascade with Raf1 immunoprecipitated from TNF-, ceramide- and sphingomyelinase-stimulated cells.

METHODS. **a**, The cells (3×10^6) were resuspended in serum-free phosphate-free RPMI containing ^{32}P -orthophosphate (6 mCi per 15 ml) for 2 h and then resuspended into the same buffer without radiolabel (5×10^7 per point) and stimulated with TNF (1 nM). At the indicated times, cells were homogenized, and Raf1 was immunoprecipitated and resolved by 7.5% SDS-PAGE as in Fig. 1. Raf1 was transferred to an Immobilon PVDF (Millipore) membrane according to the vendor's instruction. An autoradiogram was obtained and a western blot (described below) with anti-Raf1 antibody was used to monitor recovery of Raf1 protein. The data (CPM/Raf protein) are presented as fold of control and represent one of three studies performed in triplicate. **b**, HL-60 cells (30×10^6 per incubation) were stimulated with TNF (1 nM) for 20 min, lysed in RIPA buffer, and Raf1 was immunoprecipitated as above. MEK1 phosphorylation by immunoprecipitated Raf1 was as described in Fig. 1 for 15 min. Phosphorylated MEK1 was transferred to an Immobilon PVDF membrane and western analysis done with rabbit anti-MEK1 antibody (a gift from R. Davis, University of Massachusetts) using an ECL detection system (Amersham Life Science). The data represent one of three similar experiments. Identical results were obtained with cells stimulated for 5 min with TNF. **c**, These studies were done as in **b** except cells were stimulated for 20 min with C8-ceramide (25 μM) or *S. aureus* sphingomyelinase (10 Mu ml^{-1}), and MEK1 was phosphorylated for 30 min *in vitro* by immunoprecipitated Raf1. These data represent one of two similar studies. **d**, Raf1 was immunoprecipitated from cells (8×10^7 per point) and used to reconstitute the MAP kinase cascade as in Fig. 1. The last two lanes did not receive Raf1.

► 275 so that the potential CAP kinase phosphorylation site was situated in the middle rather than the carboxy terminus. Each synthetic peptide (40 μg per reaction) was phosphorylated for 30 min by CAP kinase as described in Fig. 1 and reactions were terminated with 0.5 M ATP in 90% formic acid. TFA was then added and phosphorylated peptides were resolved by reverse-phase HPLC as described in a using a linear

gradient of acetonitrile from 2–60% in 0.1% TFA at a rate of $1\% \text{ min}^{-1}$ with a flow rate of 1 ml min^{-1} . Fractions (1 ml) were collected for Cerenkov counting. The data represent one of five identical experiments. **d**, Reconstitution of the MAP kinase cascade was as in Fig. 1 using mutant FLAG/Raf1 substituted with valine for threonine at residues 268 and 269. Mutant FLAG/Raf1 was coexpressed with *src* and *ras* in Sf9 cells.

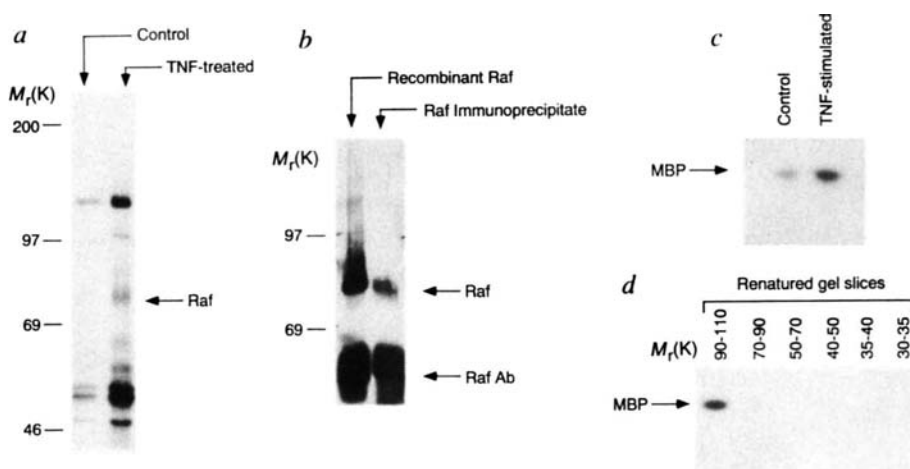


FIG. 4 Raf1 complexes with CAP kinase. *a*, Immune complex kinase assay using Raf1 immunoprecipitates from control and TNF-stimulated HL-60 cells. *b*, Western blot using anti-Raf1 antibody. *c*, A Raf1 immunoprecipitate from TNF-stimulated HL-60 cells exhibits enhanced activity towards the generic proline-directed substrate myelin basic protein (MBP). *d*, Renaturation of the 97K CAP kinase from the Raf1 immunoprecipitate.

METHODS. *a*, HL-60 cells (7×10^7 per incubation) were stimulated with TNF, and Raf1 was immunoprecipitated and incubated in Raf1 reaction buffer as in Fig. 1 containing 5 μ M ATP and 20 μ Ci [γ - 32 P]ATP. After 30 min, phosphorylated proteins were separated by 7.5% SDS-PAGE, transferred to an Immobilon PVDF membrane, and autoradiographed. These results represent one of three similar experiments. *b*, Western blot analysis using anti-Raf1 antibody was done as described in Fig. 3. *c*, Cell lysates were prepared in RIPA lysis buffer from control and

TNF-stimulated HL-60 cells and Raf1 was immunoprecipitated. The immune-complex was assayed for kinase activity towards MBP (50 μ M) in 30 mM HEPES, pH 7.4, 10 mM MgCl₂, 5 mM NaF, 50 μ M ATP, 15 μ Ci [γ - 32 P]ATP. After 20 min, the beads containing immunoprecipitated Raf1 were removed by centrifugation and phosphorylated MBP in the supernatant was resolved by 13% SDS-PAGE and autoradiographed. The data represent one of two similar experiments. *d*, Proteins contained within the Raf1 immune-complex derived from the TNF-stimulated cells in *a* were separated on a 7.5% SDS-PAGE and renatured for CAP kinase activity as described in Fig. 1. Gel slices were cut according to the chromatogram defined by the molecular mass markers and used to phosphorylate MBP for 60 min as in *c*. There was little renatured activity when the immune complex was derived from control cells. The data represent one of two similar experiments.

from TNF-stimulated cells, and was shown to correspond to CAP kinase (Fig. 4c, d).

Taken together, our results strongly implicate CAP kinase as a physiologically important Raf kinase. An obvious question is whether there is a role for *ras* protein in this signalling pathway. The *ras* protein may be involved in ceramide-mediated apoptosis through Fas²⁹, and our own preliminary studies (not shown) also implicate *ras* protein in TNF-induced MAP kinase activation.

Ceramide functions as a dual second messenger for TNF in stimulation of inflammation and apoptosis¹. Recent investigations suggest that in the latter capacity, the SAPK/JNK (c-Jun

NH₂-terminal kinase) signalling system mediates TNF action and the MAP kinase cascade is not recruited (Verheij *et al.*, manuscript submitted). Whether the response system is genetically programmed in specific cell types or whether an upstream switch determines the path through which the ceramide signal proceeds is being investigated.

Our investigations suggest that TNF receptor interaction, through ceramide generation, stimulates CAP kinase to complex with, phosphorylate and activate Raf1, linking the sphingomyelin pathway at the cell surface through to MAP kinase. □

Received 19 April; accepted 13 September 1995.

- Kolesnick, R. & Golde, D. W. *Cell* **77**, 325–328 (1994).
- Liu, J., Mathias, S., Yang, Z. & Kolesnick, R. N. *J. biol. Chem.* **269**, 3047–3052 (1994).
- Hannun, Y. A. *J. biol. Chem.* **269**, 3125–3128 (1994).
- Heller, R. A. & Kronke, M. *Cell Biol.* **126**, 5–9 (1994).
- Joseph, C. K., Byun, H.-S., Bittman, R. & Kolesnick, R. N. *J. biol. Chem.* **268**, 2002–2006 (1993).
- Belka, C. *et al.* *EMBO J.* **14**, 1156–1165 (1995).
- Raines, M. A., Kolesnick, R. N. & Golde, D. W. *J. biol. Chem.* **268**, 14572–14575 (1993).
- Vietor, I., Schwenger, P., Li, W., Schlessinger, J. & Vilcek, J. *J. biol. Chem.* **268**, 18994–18999 (1993).
- Zhang, X.-F. *et al.* *Nature* **364**, 308–313 (1993).
- Stokoe, D., MacDonald, S. G., Cadwallader, K., Symons, M. & Hancock, J. F. *Science* **264**, 1463–1467 (1994).
- Leever, S. J., Paterson, H. F. & Marshall, C. J. *Nature* **369**, 411–414 (1994).
- Daum, G., Eisenmann-Tappe, I., Fries, H.-W., Troppmair, J. & Rapp, U. R. *Trends biochem. Sci.* **19**, 474–480 (1994).
- Williams, N. G. & Roberts, T. M. *Cancer Metastasis Rev.* **13**, 105–116 (1994).
- Kyriakis, J. M., Force, T. L., Rapp, U. R., Bonventre, J. V. & Avruch, J. *J. biol. Chem.* **268**, 16009–16019 (1993).
- Kovacina, K. S. *et al.* *J. biol. Chem.* **265**, 12115–12118 (1990).

- Morrison, D. K. *et al.* *Cell* **58**, 649–657 (1989).
- Kolch, W. *et al.* *Nature* **364**, 249–252 (1993).
- Morrison, D. K., Heidecker, G., Rapp, U. R. & Copeland, T. D. *J. biol. Chem.* **268**, 17309–17316 (1993).
- Dent, P., Jelinek, T., Morrison, D. K., Weber, M. J. & Sturgill, T. W. *Science* **268**, 1902–1906 (1995).
- Macdonald, S. G. *et al.* *Molec. cell. Biol.* **13**, 6615–6220 (1993).
- Izumi, T. *et al.* *J. biol. Chem.* **266**, 7933–7939 (1991).
- Pantazis, P., Kharbanda, S., Goustin, A. S., Galanopoulos, T. & Kufe, D. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2481–2485 (1991).
- Wartmann, M. R. & Davis, R. J. *J. biol. Chem.* **269**, 6695–6701 (1994).
- Fantl, W. J. *et al.* *Nature* **371**, 612–614 (1994).
- Freed, E., Symons, M., Macdonald, S. G., McCormick, F. & Ruggieri, R. *Science* **265**, 1713–1716 (1994).
- Fu, H. *et al.* *Science* **266**, 126–129 (1994).
- App, H. *et al.* *Molec. cell. Biol.* **11**, 913–919 (1991).
- Wang, H.-G. *et al.* *Oncogene* **9**, 2751–2756 (1994).
- Guilbins, E. *et al.* *Immunity* **2**, 341–351 (1995).
- Morre, D. J. *et al.* *Meth. Enzym.* **228**, 448–450 (1994).

ACKNOWLEDGEMENTS. We thank D. Morrison for providing recombinant Raf1 lysates, Raf1 antibody and for assistance with the mapping of the phosphorylation site on Raf1, and Z. Fuks for thoughtful commentary.

Preparing for the next millennium

With scientific literacy now a necessity rather than a luxury, children's science books have taken on a new importance. James Trefil introduces *Nature's* first children's books supplement with a review of some recent reference volumes*.

THE market for children's books is one of those strange institutions, like the market for college textbooks, where the people who buy the books are not the same as the people for whom the books are intended. Parents, the purchasers, shop for science books for many reasons. The most important, presumably, is the sense that the world has changed, and that attitudes towards science that permeated the educational systems of the industrial world in the middle of the twentieth century simply aren't going to serve their children as we approach the twenty-first.

It used to be that we could get away with thinking of science as merely one of those embellishments, like a knowledge of Renaissance poetry or the Russian novel, that are the mark of an educated person but have no possible relevance to real life. That sort of attitude just won't work any more. Children growing up today are going to enter a world in which science and technology play an ever greater role in shaping everyone's life and work. To make sense of that world, to function as a citizen, that child will have to have a firm understanding of the scientific issues and concepts that will shape his or her life. We talk about this sort of intellectual framework as scientific literacy, which describes people's ability to deal with scientific issues with as much ease as they do other issues that come across their horizons.

Given this background, the first thing we have to ask in evaluating a children's science book concerns the nature of the audience for which it is intended — is it to be judged as a book that will, in and of itself, make children more scientifically literate, or is it instead intended as background for parents who will, in turn, help their child in this field? To help me with this assessment, I recruited an expert — my ten-year-old son Tomas, who represents the ultimate audience for them. Here is our consensus evaluation of a selection of general science reference books.

The Kingfisher Science Encyclopedia edited by Catherine Headlam (Kingfisher; £25, \$39.95), despite its name, is actually an expanded dictionary of scientific terms. Look up 'immune system', for example, and you'll find a half-page discussion that hits high points such as antibodies, white blood cells, AIDS and vaccination, together with a nicely done illustration of the workings of

antibodies and macrophages. The strong point of the book is the illustrations: my son and I both learned a lot just by going through and reading figure captions. The text itself may be a bit advanced for children below the age of 12 or so.

The Kingfisher Visual Encyclopedia of Science edited by Mychele Byam (Kingfisher; £17.99, \$22.95), on the other



hand, gets my vote for the most child-friendly of the publications — an evaluation that is seconded by my son's fifth-grade class. Lavishly illustrated, it is divided into four sections dealing with the Earth sciences, the living world, stars and planets, and technology (the latter section, for some reason, includes the structure of matter). Using what I think of as a "Sesame Street" format, it presents snippets of information that, taken together, present the sort of information one would expect a scientifically literate child to have. A long index can be used to locate specific information.

The Guinness Encyclopedia of Science by Robert Youngson (Guinness, £21.95), on the other hand, is organized more like a traditional encyclopaedia, with longish articles that go into specific areas in some depth, the articles being grouped according to subject matter. Under "Computing", for example, we find essays on artificial intelligence, information theory, software and other such topics. The second half of the book is a dictionary-style "Factfinder" for quick reference.

This is the book I would most recommend for parents who want to understand in some depth what their children's science classes are about. It has a straightforward, honest style, and if there is some scientific dispute on an issue, it says so without mincing words. The author is also

willing to resist modern fads. You will search in vain for deconstruction in the discussion of the philosophy of science, for example, and under a photograph of an African medicine man is the statement that people who have access to modern Western medicine are better off than those who do not — a fact that everyone knows to be true but which, on my side of the Atlantic at least, is considered unfashionable to point out.

The Oxford Children's Book of Science by Charles Taylor and Stephen Pole (Oxford University Press, £14.99) seems to me to be well suited to direct use by children in the traditional encyclopaedia mode. It is lavishly illustrated and consists of articles on specific topics such as energy, machines, probability and so on. My son felt that the illustrations made him want to read the text, and the text itself seemed accessible for someone his age. My problem with the book is that it may be difficult to dig out specific information from the longer essays (there is only a short index at the back). In addition, the emphasis in the essays seems to be heavily on traditional classical science — there is only a paragraph or two on the working of computers, for example, buried in an essay on animal communication.

Computers, whether through CD-ROMs or the Internet, are providing children with more and more scientific information and activities. Usborne's **Computer Dictionary for Beginners** by Anna Claybourne (£5.99) is a series of short essays linking together definitions of words important in information technology, together with an extensive index and a useful glossary of computer slang. Under a heading entitled "Computer networks", for example, you'll find definitions and short descriptions of LANs, modems, optical cables, passwords, servers and other kinds of terminology. There are illustrations and photos, some of which (like the cutaway view of a mouse) supply a lot of background information in a small space. The book won't get your child started with a particular computer, but it might well provide a useful reference as he or she gets into the digital world. □

James Trefil, co-author of *The Dictionary of Cultural Literacy and Science Matters* and author of *1001 Things Everyone Should Know About Science*, is in the Department of Physics, George Mason University, Fairfax, Virginia 22030, USA.

*Where possible, details of availability and price in the United States and the United Kingdom are given throughout the supplement. All readers should check prices before ordering.

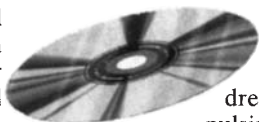
Hyper-media or media-hype?

Tim Brosnan

CHRISTMAS is coming and if you are looking for a science CD-ROM for your children you could usefully turn to **The Computer Museum Guide to the Best Software for Kids** by Cathy Miranker and Alison Elliott (HarperPerennial, \$16) for advice. This is an impartial and informative collection of reviews written by staff at the Computer Museum in Boston that summarizes more than 200 pieces of software (mostly CD-ROMs) written for children aged 2–12, giving each product between one and four stars for learning, looks and longevity.

Any package you buy should ideally entertain and enthrall your children, helping to turn them on to the joy of learning science. Such software does exist but so does much that is boring — or worse. I have looked at three pieces of software that illustrate radically different approaches to 'having fun in science' and have compared my ratings with the guide's. (One can also access the Computer Museum's list of "best buys" and regular updates on its reviews via its World Wide Web page at <http://www.tcm.org/>.)

I opened **Widget Workshop** (Maxis, \$34.95, £29.99) with keen anticipation because it is published by the company that produces the superb "Sim" range of computer games. This CD-ROM is a collection of attractive computer 'widgets' that can be connected together to perform different kinds of counting, timing and logic activities. These are imaginatively designed and the disc is accompanied by a range of "exploration tools" such as a thermometer and magnifying glass for use in activities intended to relate the computer simulations to the real world. The guide gives it two stars but notes that it is "so challenging that many kids will need your help to take advantage of everything it offers". True, but the reason I was crushingly disappointed when I used it and why I cannot give it even one star is neither this nor its penchant for f.p.s. rather than SI units, nor even that the science covered is consistently way over the level of any 8-year-old I know (its target age). What worries me most about this product is that, in an age when all involved in science education are concerned to show that science is for everyone and to destroy the stereotype of a scientist as a crazy white man in a white coat, it blandly tells us that "the most fun way to look at science is from the point of view of a mad scientist" and rewards success by letting you join the "Mad Scientist Hall of Fame, take over the world and advance to... more science



and more madness". Not only that, some of the activities are highly unethical. In the second one children are told to connect a heart icon (a pulsing switch) to a bulb with a single piece of wire to make it flash and are then told: "That's how a real mad scientist turns on a light bulb".

How many readers of *Nature* would want children to think that this is what they get up to? Is this really the best way to encourage children to believe that science is for them? I think not, and neither do the authors of **Science Adventure II** (Knowledge Adventure/Random House; £29.99, £29.99), who obviously have their hearts in the right place when they write: "the way to ignite a child's intellect is to introduce him or her to the fun of learning. If we truly achieve this, we no longer have to come up with sneaky ways to get kids to learn."

Unfortunately their CD-ROM does not achieve this. The guide gives it only one star, noting that it "fosters the random acquisition of facts" and that there is "too much looking and not enough doing". They are right. Despite containing a vast amount of information, its poor produc-

tion and lack of interaction and signposting let it down, and all who tried it soon got bored and stopped.

This was not true of the beautiful **Eye-witness Encyclopedia of Science** (Dorling Kindersley; £59, \$79.95), which captivated all who used it. Essentially it is a CD-ROM version of one of their books, but this is like describing the Grand Canyon as a hole in the ground. It has a stylish, clear structure and interface that lets you find out easily what you want to know while guiding you to paths you might not have thought of. The text, illustrations, video clips and hypertext links are attractive, thoughtful and appropriate and the slips few (an example is the widely spaced water molecules in the section on dissolving). Using it, one begins to experience the true fun of learning science with the excitement in the science not the gimmicks. How refreshing to find a hypermedia product that is more media than hype. The guide gives it a fully deserved three stars. If you like the Dorling Kindersley range of books, don't buy them — get the CD-ROMs instead. □

Tim Brosnan, the father of six-year-old twins and a nine-year-old, is in the Science and Technology Group, Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL, UK.

HANDS-ON SCIENCE

Random walks and lucky dips

Graham Farmelo

GENERALLY speaking, the standard of children's science books is pretty dreadful. Anyone who has served on the judging panel of the Rhône-Poulenc Prizes for Science Books will confirm the difficulty of finding six first-class children's books for the shortlist. The problem came to a head last year when the judges shortlisted only four books.

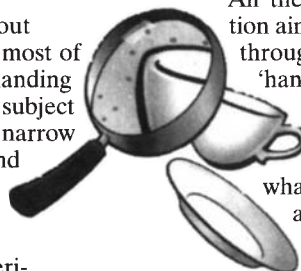
The usual complaints about these science books are that most of them are uninspired, undemanding and formulaic. True, their subject matter is limited by the narrow range of observations that lend themselves to elementary scientific discussion, by the few domestic items that are readily available for experiments and by the all-important safety considerations. Yet, even with these constraints, there is surely still plenty of scope for the talented writer and illustrator to capture the imagination of the young scientist.

In **Riddles in Your Teacup** (2nd edn; Institute of Physics; £10.95, \$22), Partha Ghose and Dipankar Home consider dozens of everyday phenomena that can

be understood using simple scientific ideas — how bees buzz, why one should not lick an ice tray, why migratory birds fly in V-shaped formations, and so on. Many will enjoy dipping into this attractive though rather feebly illustrated little book, which is likely to appeal beyond its main audience of high-school children.

All the other books in this selection aim to make science accessible through experiments and other 'hands-on' activities. Janice VanCleave's **Geometry for Every Kid** (Wiley; \$12.95, £6.95) is a clear if somewhat stolidly presented set of activities for 8–12-year-olds.

It would provide useful back-up material for students struggling with mathematics at school, although the style of presentation is so formal that few children will choose to read it of their own volition. The same could be said of the author's **Electricity** (Wiley; \$11.95, £6.95), which features 20 standard experiments about the nature of electricity. As the publisher claims, these are "simple and fun", but few children these days will find them "mind-boggling".



No-one who turns directly from Van Cleave's books to Jean Potter's **Science in Seconds for Kids** (Wiley; \$12.95, £8.99) could entirely avoid a feeling of *déjà vu*. The authors' experiments may be different, but their formats, presentational styles and illustrations are similarly dull. Potter's book does, however, have more than its fair share of unfortunate and inaccurate scientific language. What can a child be expected to understand by the definition of an atom as the "smallest part of a material that retains the properties of that material"? And is it admissible to define the viscosity of a liquid as "the rate at which the liquid can be poured"?

The San Francisco Exploratorium is justly famous for its innovative approach to hands-on science and for the development of its courses for teachers. Two books in its Science Snackbook series are **The Magic Wand and Other Bright Experiments on Light and Color** and **The Cheshire Cat and Other Eye-Popping Experiments on How We See the World** (Wiley; \$10.95, £7.99 each). They contain 26 "science snacks", miniature versions of some of the exploratorium's most popular exhibits. Originally produced for local high-school teachers, these new versions are for a wider audience. As one might expect, the experiments here are much more imaginative than those in most other science books of this type, although it has to be said that the format is depressingly predictable and not sufficiently engaging to stand a chance of appealing to the average youngster, who would much rather be doing something else.

The kitchen is the laboratory of the home and it is an ideal place to learn science, especially as one can often eat

the results of the experiments. In **The Science Chef** (Wiley; \$12.95, £9.99), Joan D'Amico and Karen Eich Drummond present "100 fun food experiments and recipes for kids". The authors have chosen the recipes well and their instructions are clear enough, but they have done too little to make the exercises scientifically rewarding. A missed opportunity.

Lucky Science (Wiley; \$10.95, £6.99) uses an unusual and attractive historical theme to whet its readers' appetites for science. Royston and Jeannie Roberts examine 20 breakthroughs that were all made serendipitously, such as the discovery of penicillin and the invention of photography. It is the same story again: good activities, shame about the presentation.

The historical theme is also used by Steve Tomecek in **Simple Attractions** (W. H. Freeman; \$9.95, £5.95), which looks briefly at the "phantastic physical phenomena" of electricity and magnetism. It has colourful, stylish illustrations which make this easily the most visually attractive of all the books reviewed in this selection: what a pity that its presentation is let down by the prolix and uninspired text.

Overall, the grade I would give this selection is D for disappointing. With the fickle attention of children being sought by so many attractive products in so many media — television, radio, magazines, films, CD-ROMs and so on — children's science books will have to do much more if they are to hold their own in the ruthlessly competitive marketplace of the edutainment industry. □

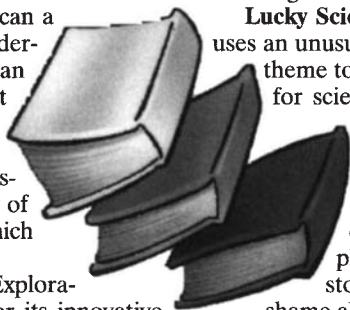
Graham Farmelo is head of programmes at the Science Museum, Exhibition Road, London SW7 2DD, UK.

the experiencing of this outwardly conventional book a truly multimedia event.

In addition to all the success of its visual and tactile aspects, something that all too often signals a neglected text in inferior books and CD-ROMs, **The Art of Construction** is also a very substantial introduction to building materials and techniques. It is well written, with a clear, accurate and engaging narrative that ranges from prehistoric structures built for shelter to examples of modern architecture built for vanity, while showing the continuity of ancient and modern ideas of building. The opening discussion of Egyptian pyramids, for example, contrasts their construction with that of the glass pyramid in the courtyard of the Louvre. There is more information included in — but not crammed into — this volume than in most books and CD-ROMs, whether for children or adults.

I Wonder Why Tunnels Are Round And Other Questions About Building by Steve Parker (Kingfisher; £6.99, \$9.95 (hbk); £3.99 (pbk)) also contains a good deal of information, but it is more conventionally presented even while being more arbitrarily arranged. As the title suggests, the book consists of a series of questions (almost three dozen) about building. The answers are generally terse, and what related and tangential information there is consists mostly of odd facts and figures. We read that "tunnels are round because drills make round holes!" as well as that "an arch is a much stronger shape than a square", but there is little elaboration beyond such statements. We read beside a bullet that if all the material excavated from the Eurotunnel were piled up it would be as tall as the Eiffel Tower, but the illustration shows a pile of dirt almost 50 per cent taller. We learn that the Sydney Harbour Bridge is nicknamed "the coathanger" and is the widest in the world, but we are not told that it is a steel arch.

Make It Work! Building by Andrew Haslam with text by David Glover (Two-Can, £7.99) is the least satisfying and most misleading of these three books. Although it is subtitled "The Hands-On Approach to Science", there is little science in this book, for it consists largely of instructions on structures to be made out of paper, wood, plastic and other common materials. The book is more about the art than the science of building and construction, and it propagates the common confusion between the fundamental objective of engineering, which is to design and build what does not occur in nature, and that of science, which is to study what already exists. Engineering science takes as its objects of study the fabricated components of buildings, bridges and



ENGINEERING SCIENCE

Building firm foundations

Henry Petroski

PRODUCING a book is a problem in design, and as such has no single solution. It should therefore come as no surprise that three books for children on the same subject can be as different as these on building and construction.

The Art of Construction (Kingfisher, £12.99; published as *Architecture and Construction* in the United States by Scholastic at \$19.95) is the kind of conventional book that can hold its own against a CD-ROM. Indeed, this book is more user-friendly, interactive and visually stimulating than most computer-based text and graphics packages that have been touted to supplant books. From the outside it looks like a conventional book, but its case-cover

opens to reveal a spiral-bound inside that allows the pages to lie flatter than a computer screen. The glossy pages are also more colourful and varied in readable typography than a digital competitor, and the use of a wide variety of textures of paper and graphical forms makes turning the 40-odd pages, some of which fold up and out, an experience of increasing anticipation as the end of the book is approached. Inside the back cover is a pocket containing a miniature version of an illustrated newspaper that employs all the best features of desktop publishing. There is also a sheet of stickers that the reader is invited to paste on pages marked with a special icon, a further way of making



the like, but words such as 'engineer' and 'engineering' appear all too infrequently in the text and not at all in the title of this book about engineering and built things. □

Henry Petroski, author of *Engineers of Dreams: Great Bridge Builders and the Spanning of America (Knopf)*, is in the Department of Civil Engineering, Duke University, Durham, North Carolina 27708-0287, USA.

EARTH SCIENCE

Getting the Earth right

Peter Nisbet

I LIKE books that are simple and not at all like school textbooks. I like science books that give you lots of interesting information. I think quite a few people think differently, but I like old facts mixed up with new, and good diagrams. My eyes are better than my parents' so I also can read small print — if it's fun, that is.

Connections! Earth by Caroline Grimshaw (Two-Can, £7.99) is divided into four sections called "Mysterious Planet", "Changing Planet", "Life on Earth" and "What Future?". This is a good book in the way that it is set out in the form of questions and answers. But many of the facts are wrong, and some of the pictures are inappropriate for the caption.

For instance, one of the questions is "How do climates divide the world?", and there is a map of the globe with each major climate represented by a different colour. In the key, one of the colours is named "Polar" but this colour does not appear on the map. The map shows Siberia and Alaska at winter temperatures of -6°C . (I didn't know that we had so much global warming!) Moving through the book, it says that the Milky Way is part of a galaxy. This is not really right, as it is a galaxy — 'galaxy' means 'milky way'. The book says the inner mantle is molten, which is not true. It also says that the inner core has a temperature of $4,500^{\circ}\text{C}$. Nowadays we think it is a bit hotter, and my mother says that it is around $6,000^{\circ}\text{C}$ in the middle. I don't much like the strange typeface and the way that some text is on coloured patterned paper. The index is especially hard to read because of this. The front page has a nice satellite picture of a city, but it doesn't say that it is Washington DC.

Earth and Space by Susan Mayes and Sophy Tahta (Usborne; £7.99, \$17.95 (hbk); £5.99 (pbk)) is also divided into sections but there are six of them. They are "What's the Earth made of?", "What's under the ground?", "What's under the sea?", "What makes it rain?", "Why is night dark?" and "What's out in space?". Parts of "What makes it rain?" are a bit too

simple, but other parts are really good. As well as having very few errors, this book is five times as long as *Connections! Earth*. Like all my other Usborne books it has good facts and good colour pictures. It is not quite right when it says that there is hot sticky liquid rock in the mantle that churns around and makes the plates move. Otherwise it's super. It is rather strange, though, because its content is really for 10–12-year-olds but the size of the writing is meant for younger people. But I think that younger children would enjoy it too. Did you know that Antarctic cod has special chemicals in its blood to stop it freezing? Or that the longest tunnel is almost 169 km long and carries water to New York City?

If I had to choose between these two books for myself or the school library I would choose *Earth and Space*.

Peter Nisbet (aged 10) is at St Jude's Church of England School, Englefield Green, Surrey, UK.

QUANTUM PHYSICS

Curiouser and curiouser

Alfred Mallet and Edward A. Knapp

ALFRED: Russell Stannard's *Uncle Albert and the Quantum Quest* (Faber, £3.99) is about a girl called Gedanken and some of her adventures with her eccentric Uncle Albert. She goes to visit him a lot and he can make a thought bubble, and if he thinks very hard, she goes into the bubble. Children will enjoy the book because the story is fun and exciting, but it is too complicated and densely packed with some of the most interesting physics. The White Rabbit and nuclear raspberries make it fun, but Gedanken's confusion leaves the reader confused.

ED: Although Uncle Albert seems to be correct in the description of quantum theory, to a young reader this just has to be extremely perplexing and very *ad hoc*. After all, quantum theory did emerge from a set of puzzling experimental results that could not be described by the classical physics of the day. The particle-wave duality is described in the book in terms of diffraction — a phenomenon with which no youngster could have had experience. [**ALFRED:** How does diffraction happen anyway?] The book is an enjoyable read for a physicist, but we doubt whether any child will emerge with more real understanding of quantum theory than when they began the quest. □

Alfred Mallet (aged 8) is at The Oratory RC Primary School, Chelsea, London, UK; Edward A. Knapp is Alfred's grandfather and is president of the Santa Fe Institute, Santa Fe, New Mexico 87501, USA.

GEOGRAPHY

Global expeditions

Anne Surridge



3D Atlas: A Multi-Media Expedition to Understanding Planet Earth (Creative Wonders; \$49.95, £55) can show lots of different globes, such as an environmental globe, a physical globe and a political globe. You can turn these to see the world from different angles and you can zoom in to look at different countries — England can be made to fill the whole screen. It has 3D flight simulations over places like the Rockies and short narrated stories about things like global warming and volcanoes. There are other globes that show the world at night and the Earth's crust. It also has a game called "Around the World", with lots of challenging questions.

The best features are the colourful graphics. There are good movie clips with the stories and in the 3D flights you really believe you are flying in an aeroplane. The sound and narration are also very good. There is lots of information about the countries: the atlas not only shows you where in the world things like countries, cities and rivers are but it gives you facts and pictures as well. I found the stories useful for finding out about the ozone layer and the sea level too. The atlas has already helped me with a school project — there is so much in it that you learn more and more every time you use it.

I did not find many problems with the **3D Atlas**. It was very slow on my machine (486 DX2 with a 66 Mhz clock) and it can sometimes be confusing as there are so many things it can do. It sometimes doesn't do what you want it to. For instance, you have to be careful when rotating the globes or you can turn too far, although this might be easier to avoid with a faster computer. The title list is also a bit erratic. Some of the features are a bit disappointing, such as the world at night which I expected to look prettier.

The **3D Atlas** is easy to use. Everything is normally straightforward and quite simple. If you really get stuck you can go to the help section, which gives good advice. The game always makes it clear what you have to do — I did not need to use the instruction book at all.

I liked the **3D Atlas** a lot. I think anyone would find it interesting and fun, and it will improve their geography too. I learnt many things and it is definitely useful for homework. It has a few problems, but I think it is really good and exciting. You probably need a fairly fast computer for it to work well, so if you have a computer like mine you need to be patient.

Anne Surridge (aged 11) is at Sedgemoor School, Catford, London, UK.



Depicting the facts of life

Sandra Knapp

NATURAL history has always fascinated children, and children are naturals as natural historians. What adult has not gone for a walk with a child and had to drag him or her away from minute examination of some insect or worm? But it is worrying what a limited scope of topics publishers seem to think will attract children and their parents (whose opinions on what constitutes a good book in turn also differ rather alarmingly). Predictably, vertebrates and ecology feature prominently in the children's books reviewed here, and if coverage reflects the relative densities of animals on Earth, then we are overrun with mammals, and with top predators at that. Children are assumed to be unable to enjoy learning scientific names of any organisms, except of course dinosaurs. The euphony of *Tyrannosaurus rex* enchants children of all ages, but why not also *Bactris gasipaes* or *Homo sapiens* or *Morpho eugenia*?

For enjoyment together

Stories about nature are a painless way for a child to be introduced to the science of natural history and for adults to participate in the learning process. Joanna Foster's **The Magpie's Nest** (Clarion, \$15.95), a beautifully illustrated re-telling of an old folk tale about why birds have different types of nests, is sure to please children and parents alike. Kathryn Lasky's **Pond Year** (Walker, £8.99), the story of two six-year-old girls' annual cycle of play in and around a New England pond, is some of the best writing for children I have seen, and ranks high among story books as well as among books about nature. The gorgeously surreal illustrations in Bert Kitchen's **When Hunger Calls** (Walker/Candlewick; £4.99, \$15.95), combined with good solid information about animal's eating habits, make this an ideal children-and-parents' book. My favourite is the Galápagos woodpecker finch in a forest of upside-down *Opuntia* cacti, with an understorey of what actually seems to be *Cannabis*, whereas my four-year-old's is the killer whale attacking the sea lion on the beach.

Pop-up books can be a good introduction to nature. **Busy Beaver Pond** and **Lion Cubs at Home** (W. H. Freeman; \$8.95, £5.95) in the One Very Small Square series are terrific and provide good basic information with questions for a child to ponder. Information is the strong point of **My First Pop-up Book of Prehistoric Animals** (Simon and Schuster, \$12.95); it is a pity about the cartoon-like illustrations, but children don't seem to mind, and are thrilled by trying to pronounce *Brontotherium* and *Macrauchenia*. **Nature By The**

Numbers With Pop-up Surprises (Simon and Schuster, \$12.95) is a counting book on the theme of animals, with nice illustrations and some thoughtful talking points about natural history. The principle of camouflage never fails to fascinate children, and the 'pull-the-tab' book **Animals Cover Up** (Simon and Schuster, \$10.95) is another good example of the pop-up format used to excellent effect.

For slightly older children already keen on observation, the books by Karen Bryant-Mole in the See For Yourself series (**Soil, Trees, Flowers** and **Insects**) are excellent (A&C Black, £6.99 each). They fit well into topic work in the UK National Curriculum, and pose questions children can answer by observation. An adult is necessary, though, because some of the photographs need interpretation and gentle assistance (for example, the ovules in the daffodil in *Flowers* are hard to see unless you already know where to look). Although the photographs are wonderful and certain to spark children's imaginations and observational skills, these books somehow do not escape being rather dull and textbookish.

For stimulating interest

Children are natural observers, but channelling their thoughts with directed tasks can greatly enhance their enjoyment of the process. **Nature in a Nutshell for Kids** by Jean Potter (Wiley; \$12.95, £8.99) is just the book to help with this. It is a collection of short, easy investigations, on topics ranging from "Why are Rocks in Rivers and Streams Smooth?" to "What Animals Make Sounds at Night?", all of which can be done with minimal equipment. The book is organized by season and each investigation is framed as a question — a brilliant way to introduce hypothesis testing. Some of the tasks will require adult help and interpretation, but most can be done by children on their own. This is an ideal book for budding scientists of all descriptions, not just natural historians.

Two new additions to the Dorling Kindersley Eyewitness series, **Gorilla** by Ian Redmond and **Arctic and Antarctic** by Barbara Taylor (£8.99, \$12.95 each), are predictably lavish and interesting. They have a few glaring errors (a lichen labelled as a moss and a distinct dicot said to be a grass) but these do not detract from their overall successful format. These books appeal to children of all ages — those who

cannot yet read are quite happy to spend hours just looking at the illustrations. *Gorilla* is oddly named, because it actually is about primates and their relatives. The relationship of humans to the primates is not skirted, but is presented as fact: the first page is a comparison of the skeletons of the four primates. This is not so in the other book on primates, **The Mountain Gorilla** by Melissa Kimm (Riverswift, £4.50). Here great apes are



said to be the "closest animal relatives" of humans — technically true, but avoiding the obvious conclusion that humans are apes. The book concentrates, however, on conservation issues, and does so sympathetically while introducing some worthwhile natural history along the way.

From Usborne come **Life on Earth** (£7.99, \$17.95 (hbk); £5.99 (pbk)) and Caroline Young's **The Great Animal Search** (£8.99 (hbk); £5.99, \$9.95 (pbk)). The Starting Point Science series of books, of which *Life on Earth* is a compendium, are excellent: the cleverly packaged bits of information in a small paperback format appeal greatly to children. The re-packaging of six previously published volumes (*Where Did Dinosaurs Go?*, *How Does A Bird Fly?*, *How Do Animals Talk?*, *Why Do Tigers Have Stripes?*, *What Makes A Flower Grow?* and *How Do Bees Make Honey?*) into *Life on Earth* is less successful. They all have different designers and authors, and so sit together somewhat uneasily. Although an index is provided, information can be frustratingly difficult for children to find and sometimes there is not enough: *How Do Bees Make Honey?* should have more detail than: "In the honey sac, the nectar gets thicker and turns to honey".

The Great Animal Search is by far and away my four-year-old's favourite book. This is a spectacularly illustrated puzzle book that challenges children to spot animals in incredibly crowded habitats. To be fair, the author says up front that in real life not so many animals would be together all at once, but the emphasis on mammals and top predators is annoying. Some of the animals are tricky to find without a grown-up's help. Nevertheless, this is a definite

kid-pleaser, with useful information thrown in as well.

Richard Orr's Nature Cross Sections (Dorling Kindersley; £12.99, \$17.95) has a similar format, but lacks the puzzle element. Here crowded habitats are pulled apart, so a child can see inside. The illustrations are exquisite, but the fold-out sections of the book will ensure that its life in an active family is short. Even I found the page on the Arctic difficult to open out. The book needs to be enjoyed as a family; in ours, this sort of book is known as a "gently, gently" type. On an exciting Christmas morning, I predict that it will last about ten minutes before being damaged beyond repair. The inclusion of marine and riverine habitats is one of the great beauties of this book; they are too often left out in favour of the rainforest.

The Living Forest (Kingfisher, £12.99; published as *Trees and Forests* in the United States by Scholastic at \$19.95) is another such fragile book. Fold-out windows, see-through pages and stickers all conspire to capture a child's imagination and interest. An adult is needed to open out the tear-off flaps and set up the windows for a child, as the book does not come ready to go. The text is well written and tremendously detailed. It is refreshing to see humans treated as part of nature, a feature conspicuously lacking in many other children's books about natural history. The book may not last long, but the whole family will enjoy it immensely.

Downs Matthews's **Wetlands** (Simon, \$15) is specifically about the southeastern United States, but will have universal appeal thanks to its lovely photographs and readable text. This is a book for children that does not treat them like idiots with the attention span of a microbe. The habitat as a whole is presented, rather than just its more sensational aspects. This book could easily be read to a very young child or used by an older one to glean information for a school project.

William Lynn Baker's **Dancing Honeybees and Other Natural Wonders of Science** (Contemporary Books, \$9.95) is not just a book about natural history; it is also a fabulous compendium of fun facts about the science of the world around us. Topics as diverse as oil spills, bats, fire ants and the extinction of native species owing to the introduction of exotics are dealt with in double-page spreads with illustrations and, uniquely, references to sources. If I had to keep only one of the books sent for review, this would be it. The children also loved it.

For delving deeper

Books for older children tend to be either compendia of facts or series of experiments. The extent to which experiment

and observation are divorced in books for this age group is alarming. Science is about fitting observations to models: Charles Darwin's discovery of evolution by natural selection — surely one of the greatest advances of modern science — relied mainly on observations of natural history rather than primarily on experiments.

Plant Biology Science Projects by David R. Hershey and **Animal Behavior Science Projects** by Nancy Woodward Cain (Wiley; \$14.95, £9.99 each) are designed for science fairs, a US institution where children devise and do experiments and compete for prizes with their results. Books that provide children with projects seem to me to defeat the purpose. An experienced judge of science fairs tells me that to win, all a child needs to do is get some original data and plot it on a graph, and that the more imaginative the project, the better its chance of winning. The projects in *Animal Behavior*, however, would be enjoyable for a family or school group and will teach children ways of handling and presenting data.

The Usborne Illustrated Encyclopedia of the Natural World (Usborne; £12.99, \$16.95 (hbk); £9.99 (pbk)) is disappointing. It is better described as an encyclopaedia of the mammalian world: a third of it is devoted to mammals. This is not to say that mammals are not important, but a page ought also to have been devoted to beetles. The book is riddled with errors and gross oversimplifications. The classification at the back is laughable: relationships among the chordates are much more resolved than those depicted, and monocotyledons certainly do not all have narrow leaves — palm trees and banana plants are both monocotyledons. Children deserve better.

The remarkable parallels between the products of natural selection and human inventions are ingeniously presented in Phil Gates's **Wild Technology** (Kingfisher, £12.95; published in the United States as *Nature Got There First* at \$17.95). Along with the comparisons of roses to razor wire, squid to jet engines, and tropical trees to cathedrals (to name just a few), children are introduced to some of the scientific principles that make the designs work. It is refreshing to see a book in which the child is respected enough to be given information at several levels. The illustrations are eye-catching and slightly bizarre (a half albatross/half-glider springs to mind) and beautifully complement the text. A slight

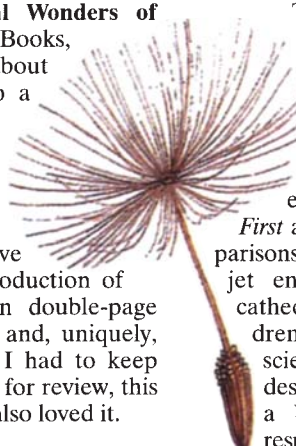
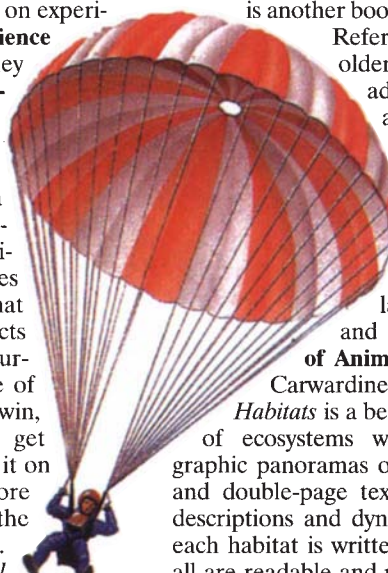
drawback is the use of different fonts to indicate comparisons: many children will not have cottoned on to this adult design trick, but a bit of explanation should clear up any confusion. Children will spend hours marvelling over the examples, and are certain to find some of their own. This is another book to look out for.

Reference volumes for older children (and adults) are often dry and boring — not the sort of thing to read in front of the fire in winter. Two exceptions are **Habitats** edited by Tony Hare (Macmillan, New York, \$25) and **The Guinness Book of Animal Records** by Mark Carwardine (Guinness, £14.99).

Habitats is a beautiful encyclopaedia of ecosystems with fold-out photographic panoramas of plants and animals and double-page text spreads of habitat descriptions and dynamics. The essay for each habitat is written by a specialist and all are readable and packed with information. But one contributor tells me to beware: he says that his piece is badly sub-edited and that errors have crept in. This is definitely a book for the secondary-school end of the market, or for exceptionally keen younger children. *The Guinness Book of Animal Records* is great fun and an excellent reference to boot. Children are fascinated by strange facts, and this book is full of them. Did you know that one species of earwig lives only in the breast pouch of a bat or that the incisors of vicuña grow continuously? Well, I didn't either. It was only with great difficulty that I prised this book away from my eight-year-old. Most of the book is devoted to vertebrates, but even microbes get a look-in. Unusually, and I think it is a great plus for this book, scientific names are used throughout at all taxonomic levels from family to species. But the book is not only about the biggest and smallest; it also contains a lot of sound information about relationships of groups of organisms, all based on recent evidence.

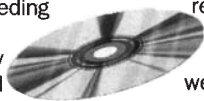
Seemingly similar, but not in the same class, is the **Nature Quiz Book** (Collins, £3.99). "Trivial Pursuit" buffs will enjoy its style and content, but it is not a learning tool, nor is it science. Also in the realm of non-science is **The Natural Hedgehog** by Lenni Sykes with Jane Durrant (Gaia Books, £8.99). This is essentially a manual for the care of sick and injured hedgehogs with a bit of natural history thrown in. It is not a children's book; rather, it is for adults whose children (or lack of them?) leave them enough time to take up aromatherapy and homeopathy for wild animals.

Children today are fortunate in having available to them a host of beautifully



Digital nature

MULTIMEDIA systems provide the opportunity for children to hear sounds of habitats and animals they will never experience. Dorling Kindersley's **Eyewitness Encyclopedia for Nature** (£59, £49.95), an "essential multimedia reference guide to the natural world", is sumptuously produced and packaged, but curiously unsatisfying. It scorned my CD-ROM drive for not having a sound card, despite not really needing one to work. The sound is a nice addition, but why does the slightly annoying American accented (and I am an American!) voice-over tediously read the simple text yet remain silent when complex Latin names appear? The small errors similar to those in other Eyewitness publications on which the CD-ROM is based will be noticed only by someone with prior knowledge of the subject, but are nevertheless irritating. It would have been easy to get it right for the CD-ROM package and very much worth the effort — so why wasn't it done? The habitat windows and video footage are more fun, and here sound is in fact useful: a child can hear a scarlet macaw call, learn about leaf-cutter ants or just listen to



the sounds of the Australian bush. Getting around each window is slow, and will frustrate all but the most computer-fanatical child.

How Animals Move (Maris; \$49.95, £39.99) is an example of just how good CD-ROM software can be. It is clearly designed as a learning tool rather than an educational computer game. Each movement type is accompanied by video footage, text and hypertext cross-referencing. There are games, but they all teach a child about the presentation of scientific data as well as the principles of animal movement. The text by R. McNeill Alexander, a world authority on biomechanics, has a depth and clarity that only a real expert can bring to a subject. Scientists can write for children, and when they do, the product is often worthwhile and exciting. *How Animals Move* is apparently designed for older, perhaps secondary school or early undergraduate, students, but the sight of two eight-year-olds discussing the effects of viscosity and drag while using it brings home its universal appeal and amazing teaching quality. The standard set by this software will be hard to beat. **S. K.**

packaged and presented books about the natural world. Meanwhile, scientists working with whole organisms and in the disciplines of taxonomy, systematics and to a certain extent evolution are becoming fewer and fewer. Could part of the problem be the persistent message of pretty pictures and lots of facts, but little science, that is so often put across by chil-

dren's books about nature? If so, perhaps we natural scientists need to get out there and write the books (and CDs) to engage and stimulate our own replacements. □

Sandra Knapp, the mother of a four-, a six- and an eight-year-old, is in the Department of Botany, Natural History Museum, Cromwell Road, London SW7 5BD, UK.

PALAEONTOLOGY

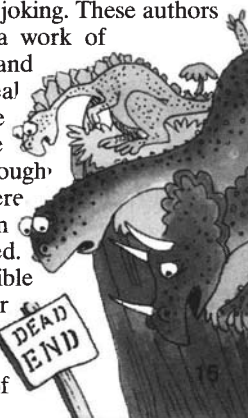
Fossils for the next generation

Douglas Palmer

I AM not sure whether today's computer kids still read books, let alone those about fossils other than dinosaurs. But one thing I am sure about is that they expect good graphics. I have tested a small selection of new books on two children, 12-year-old Amy and 13-year-old Ben.

Two general palaeontological texts on 'Life, the Universe and Almost Everything from Planet Earth to Parental Care' are stylistically as different as one can get. With a title like **Planet Ocean: A Story of Life, the Sea, and Dancing to the Fossil Record** (Ten Speed Press, \$29.95), one can guess where the authors Brad Matsen and Ray Troll originate from. The text and cartoon-style paintings confirm one's expectations. Never mind what the kids think, I like its blast of slightly wacky, Brautigan-style post-hippy, up-beat American enthusiasm for the story of fossils, the palaeontologists who find them and their

ideas. The fossil record comes across, not unreasonably, as one big lucky dip. After all, it is difficult not to wonder, sometimes with awe but more often than not with amusement, at some of the organisms that turn up: look at *Hallucigenia* and *Wiwaxia*. Perhaps God was joking. These authors are not: this is a work of deep interest and concern by real amateurs in the best sense of the word. Amy though, the pictures were "sweet" but Ben was unimpressed. The text is accessible to some extent for 9-year-olds but be prepared to answer a lot of questions.



David Norman's **Prehistoric Life: The Rise of Vertebrates** (Boxtree, £14.99) is much more strait-laced. This is a generally well-informed text, detailed and up to date. Ben was greatly taken by John Sibbick's fine paintings and drawings ("Look at the detail — how does he do it?"), as well as by the numerous photographs. As a sequel to the author's *Dinosaur!*, this is a useful volume and I particularly like the questioning and open-minded approach. So much of palaeontology is like an under-baked loaf: in need of proving but with bits that are overdone. The book is suitable for teenagers but younger 'dino-mites' may need some help.

The topic of **The First Humans** by Herbert Thomas (Thames and Hudson, £6.95) has in the past been one of those overdone bits of palaeontology, where impressive houses of cards have been built on an insecure foundations. And, as readers of *Nature* know well, important new discoveries and ideas on early human evolution appear every few months. But the predominantly historical approach taken by this little book means that most of it is up to date. The historical illustrations are especially appealing, ranging from engravings of Genesis from the 1732 Scheuchzer Bible to grainy photographs of the Anti-Evolution League in mid-1920s Tennessee and the notorious Scopes trial. This book, like others in the New Horizons series, is, I suspect, aimed more at interested adults than at children.

Not so Deborah Heiligman's **Mary Leakey** (W. H. Freeman, \$5.95, £2.95). Amy devoured this biography in one go and enthusiastically related Leakey's remarkable life story practically verbatim — although I wonder whether Leakey's success as a putative role model has as much to do with her reported hatred of school as with her fascination with human ancestry. With very traditional charcoal drawings, the book is, like a good folk tale, full of difficulties overcome. Clearly a hit, even though it is primarily aimed at 8–11-year-olds.

Meanwhile, back on the range in Texas, Louis Jacobs's **Lone Star Dinosaurs** (Texas A&M University Press, \$27.95) tells a homely American family tale of hard work and togetherness in the search for and discovery of dinosaurs — not the experience of Mary Leakey's youngest son Philip, who, according to Heiligman, complained "all I get at home is stones and bones". Nevertheless, the potential thrill of finding fossil dinosaurs or hominids certainly appeals to my young readers, but being British they know they have little hope of realizing it: "Where would we get the money from?" □

Douglas Palmer, a freelance science writer, is at 31 Mawson Road, Cambridge CB1 2DZ, UK.

The body in question

Alan Maryon Davis

YOU'D think it would be easy for me, but being a doctor actually makes it harder. I spend hours poring over the human body books in the children's section, trying to find something that will grip the imagination of my two daughters, aged 10 and 12, but in the end I do the sensible thing and let them choose for themselves.

They immediately warm to a book that, in the 20 years since it was first published, has become a minor classic. Now fully revised and updated, Judy Hindley and



Colin King's **How Your Body Works** (Usborne; £7.99 (hbk); £4.99, \$7.95 (pbk)) is a lively romp through the mechanics of the body — or, rather, a wonderfully imaginative 'Heath-Robinson' version of how it might work. The cartoon pictures are all levers, gears, conveyor-belts, buckets, hods, funnels and telephone switchboards, with little people pulling, pushing and shoving this or that. It is a riot of busy internal goings-on, and guaranteed to intrigue any child. It begins brilliantly with the digestive system — great scope here of course, with a suitably yukky chunk on being sick — and goes on to cover all the main bodily functions. The text is simple and clear for children of about 8–12 years old, with frequent questions such as "Why do you have pains?" and "What is a scab?". "Great fun", says Lizzie, my 10-year-old.

Although *How Your Body Works* is all physiology and very little anatomy, **Make It Work! Body** by Andrew Haslam with text by Liz Wyse (Two-Can, £8.99) is the opposite. It sets itself a difficult task — to convey a knowledge of the main organs and parts of the body by getting children to make models of them — but I don't think it quite succeeds. The book is liberally illustrated with colour photographs showing step-by-step how to put the models together out of common-or-garden materials. But not only are most of the models fairly fiddly to make, but in a way they also distract from learning about the function. There is little description of this, because most of the space is taken up with design and technology. Imagine spending a morning carving the atlanto-axial joint out of polystyrene, and then finding it doesn't do much to explain why we can turn our heads in all directions. Okay for the occasional school project, but not particularly educational. "All right if you like getting in a mess", says Lizzie.

Moving to a slightly older age group, we come to **The Egg and Sperm Race** by Fran Balkwill and Mic Rolph (Collins, £12.99) — the same stable that produced the award-winning *Cells Are Us* and *Cell Wars* — beautifully illustrated by Rolph with gorgeous pictures of cells, organs, double helixes and the panoply of human biology. My only disappointment is the lack of labels. Is it a brain or a heart or a kidney? Yes, I know the answer, but a lot of children won't. Balkwill's text is lively and informative, set out in undaunting short blocks. The most inviting sections are those in the form of illustrated strips of words and pictures, and there is a fascinating section at the end looking at what your body has done in the past 60 seconds —

made about 300 million red blood cells, grown one ten-thousandth of a millimetre of toenail and so on. "Yes, I like this one", says 12-year-old Jessie, "well set out with very good pictures".

And for teenagers there is a new book from Balkwill and Rolph, with the help of Victor Darley-Usmar, **Microbes, Bugs and Wonder Drugs** (Portland, £12.99; distributed in the United States by Cold Spring Harbor Laboratory Press at \$20). This is essentially a book about drugs — therapeutic and 'leisure' drugs — with potted histories of their discovery, the way they work and the use they are put to. The subject is

covered in surprising depth — the importance of the tertiary structure of drug molecules and receptors for instance — strong stuff for 14-year-olds. The story of AIDS and HIV is there. So too is herpes simplex and acyclovir. And of course we have penicillin, aspirin, anaesthesia and addictive drugs. Once again, the text is written in a clear accessible style, although it comes in rather off-putting long blocks. The pictures are as good as ever, and work best when linked together in strip form. "Really good stories — I learnt a lot", says Jessie.

Finally a grown-ups' book that is a treat for older children as well: **The Guinness Encyclopedia of the Human Being** by Robert Youngson (Guinness, £21.95). This is a most handsomely turned-out work of reference: a large-format hardback, with masses of information and profusely illustrated. It starts with how we evolved and then moves from body system to body system, taking in a smattering of psychology and human behaviour on the way. There are little snippets of history, and glimpses of up-to-the-minute technology, and the whole is thoroughly cross-referenced to guide the reader around. I particularly welcomed the useful "Factfinder" feature at the back — a combined index, medical and biographical dictionary. "All a bit too much for me", observes Jessie. "Maybe when I'm a bit older". Hmmm — maybe. □

Alan Maryon Davis, a broadcaster and writer on health matters (including children's books), is in the Department of Public Health Medicine, United Medical and Dental Schools of Guy's and St Thomas's Hospitals, London SE1 7EH, UK.

Making sense

Fran Balkwill

BEDTIME story. Mother-by-night, scientist-by-day picks up a simple 'body book' and reads: "Cells are blobs of grey jelly with a hard bit called a nucleus in the middle" — not an entirely satisfactory introduction to cell biology. But how to write a better one?

First, find a topic of universal appeal to children that can be reduced to a minimum of words and a maximum of pictures. The challenge (and ultimate reward) of distilling essential concepts into a language that children can understand lies in deciding what to leave out and in maintaining accuracy without complicating things. Young readers need to assimilate one point before proceeding to the next; they gain confidence by grasping a few basic concepts that they can return to again and again.

Science lends itself to storytelling, and a strong narrative style combined with a beguiling and inventive visual language (rather than the turgid abstract phrases of scientific journals) will help to maintain a child's interest. Used sparingly, scientific terms can be popular with children, especially if accompanied by phonetic spellings, such as 'dee-oxy-rye-

bow-new-clay-ick acid' (with apologies to American speakers).

Print runs are usually larger than those for academic books. In Britain, for example, more than 10,000 copies can be sold if the title is adopted by high-street bookshops and book clubs. And it pays to remember that children's science books have international appeal. With foreign editions, sales may exceed 100,000 copies, particularly if the book is part of a series: many publishers and bookshops are enticed by four or more volumes with a common theme or design.

Fun and humour are important: science can be daunting, and a reverential approach may alienate youngsters. That said, anthropomorphism is best avoided: 'Sid the Satellite' is invariably impossible to sustain or contain. But the key to a successful children's science book is the happy marriage of text with illustrations. For it is pictures that initially lure a child in. □

Fran Balkwill, author (with the designer and illustrator Mic Rolph) of several science books for children, is at the Biological Therapies Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.

A guide to what's out there

Barry Madore

My two children could type before they could write; and now they can 'surf the net' with the best of them. After all, they are surrounded in their daily lives by terminals, keyboards, television sets and videos. My wife (Wendy Freedman) and I are both astronomers, and we spend a large part of our lives reading and writing e-mail. But our two children still read the printed word as if books were going out of style. And when they settle down to go to sleep they are not curled up around a warm disk drive, nor do we ever see the faint glow of an active-matrix colour display under their bedcovers. Rather, they listen to the ancient words of Homer recounting an odyssey embarked on centuries ago, and hear the modern words of Lynne Reid Banks reliving contemporary stories of fantasy, adventure and strife. Blurry-eyed, they strain to see the artist's rendering of a blinded Cyclops. Fading into sleep, they see colour images of Jupiter, another one-eyed giant, this time a planet, not the son of a god, pierced by a cometary spike as accidentally tossed about the Solar System as Ulysses and his men were tossed about the Mediterranean Sea following the siege of Troy.

It is books that bring these images and words to children's minds, conveying the joy of discovery, the fascination of enquiry and the unfolding realization of what potential lies ahead in their lives. If the path from inquisitive child to research astronomer can be told, then books must surely play a central role. So what influence, if any, will the half-dozen recently published children's books on astronomy have on the current generation of youngsters? Each claims in one way or another to "convey the excitement of scientific discovery... in a lively and accessible way" or to allow readers to "experience the magic of the Universe" or "join an amazing tour through space". Needless to say, some of them fall short of their stated goals; happily, though, one or two are right on the mark.

Colin Ronan's *The Universe Explained* (Marshall Editions/Holt; £16.99, \$35) should not be judged by its plain cover. Instead, open the book to any page and let the superb images wash

over you. Skip randomly to any other page, read a few lines or a caption or two and continue. Remarkably, this is exactly how this creative book is designed to be used. Although there is an introduction, there is no real beginning to the material covered and no end to the ways in which one can read it. The book evokes the exploratory nature of astronomy, a science of the infinitely large inextricably intertwined with the infinitely small. It represents discovering astronomy at its best: highly visual, colourful yet informative, with enough text to entice a young mind without numbing it. Every two-page spread is self-contained yet cleverly tied to a dozen or more other topics by cross-references in the margin, taking the reader from "The Colour of Stars" through "Putting Light to Work" to "Echoes of the Big Bang" and so on.

Make It Work! Space by Andrew Haslam with text by David Glover (Two-Can, £8.99) did not work for me. A "hands-on approach" to the infinity of space is a tall order, and, with astronomy traditionally being an observational rather than experimental science, it is all but impossible to construct experiments for children that portray the discipline accurately. Ball-bearings in a pan of flour look nothing like any meteor craters I have seen — splashing water into a tray of mud would be a more realistic analogy. And the printed instructions are often so terse as to be useless. We are told that "each [model] planet should be made from two identical card discs... Draw rings to show the different layers within each planet... Fold the sections and slide them together to assemble your planets." Did I learn anything about planets from this clipped exercise in origami? No. Nor, I imagine, would any child of any age or sophistication. And that lesson sums up this slender hardback. It is concerned more with paper, plastic and string than it is with space or astronomy. And if you do not have a spherical sponge, a bradawl, luminous paint, a ballpen hammer, a jigsaw and a pair of 8 × 50 binoculars lying around

the house then you are apt to be just a tiny bit frustrated in trying to do many of the projects. More serious, though, is the lack of science in the book. There is regrettably little to learn from reading the text and even less to inspire a child to pursue science further or contemplate ever becoming an astronomer.

Despite differences in size and cost, the next two books invite comparison. *Astronomy* by Stuart Atkinson (Usborne; £6.99 (hbk); £4.99, \$7.95 (pbk)) is an inexpensive 'formula book', meaning not that it is crammed with equations but rather that, following other books in the series, it has a distinctive style: punchy single paragraphs under bold-type banners such as "Why is Mars Red?", "Amazing Discoveries" and "Venus Revealed". Although professing to be about astronomy, the book in fact focuses on the Solar System, with the rest of the Universe getting four out of the 48 pages. But it is certainly up to date: it is the only book in the batch that has anything on the nature and existence of the trans-Plutonian minor planets Karla and 1992~QB1. For 10–14-year-olds with short attention spans, this book is a good sampler; and if they like this one, then they will probably also like the other books in the series.

The Official London Planetarium Book of Space by Sue Becklake (Virgin, £12.99) is aimed at a similar age group, and although it contains "specially commissioned artwork", the format is reminiscent of a scrapbook, with pictures pasted helter-skelter, squares and circles, black type on red background on one page, purple type on an orange background on the next — change for its own sake without any coherence or useful effect.

The delightful three-dimensional star charts in *Reaching for the Stars* (Kingfisher, £12.99; published as *Exploring Space* in the United States by Scholastic at \$19.95) will alone entice astronomy enthusiasts of all ages, although it is difficult to say what audience the book as a whole is aimed at (9–12-year-olds would be my best guess). The editors have tried hard to produce a fun hands-on book that is useful, gripping and informative, yet the material is curiously eclectic in both its vignettes of people (Lord Rosse is highlighted, but Charles Messier goes unmentioned) and its interesting "Time Line" of space (Yuri Gagarin's name appears, but not Neil Armstrong's), literature (Patrick White's *The Tree of Man* next to Shakespeare's *Romeo and Juliet*) and visual arts (Bridget Riley's "Current" paralleled with Michelangelo's "The Last Judgment" in the Sistine Chapel). It claims to be a "stunning toolbox of special effects", and I was indeed enchanted by every page, but the spell kept being broken by the inadequate text.

I was often left hoping for more, but was unable to find it.

Russell Stannard's **Our Universe** (Kingfisher; £9.99, \$14.95) satisfied my hunger. Here is a delightful book, written as a dialogue between the author and his young readers, weaving questions and answers into a story that is factual without being dry, exciting without being over-sold. What do you need to read this book? "Imagination", replies Stannard. "There is no other way to make the journey."

Philip's Atlas of the Universe by Patrick Moore (Philips, £25) is a serious book. It is hugely informative and stamped throughout with its author's hallmark: engaging prose, a blistering pace and plenty of the latest colourful photographs from space missions and amateurs, beautifully executed original drawings and information-packed tables. I was once told that Moore periodically reserves a hotel room for a week, packs up all the illustrations for his latest book, moves in, shuts the door and starts dictating text as fast as his secretary can write it down. Be that as it may, Moore knows astronomy as does no other popularizer and he can write about it in a way that is accessible to all educated interested readers. But as its title implies, this is no introduction to the Universe; rather, it is an atlas that belongs in the classroom. And every serious armchair astronomer (of any age) will welcome it as a reference book that can be opened to any page and read with pleasure.

To end on a similarly high note we have David Levy's **Skywatching** (Harper-Collins, £14.99). Levy is an amateur astronomer in the best sense of the word: not only does he find comets, but, as shown by last year's collision of comet Shoemaker-Levy 9 with Jupiter, he is also capable of losing them. His experiences and love for the subject come across vividly in his writing. This "Ultimate Guide to the Universe", ostensibly aimed at those brave souls willing to forgo late-night television for a real show in the night sky, has been masterfully produced in a tradition reminiscent of Fred Hoyle's *Astronomy* and the many editions of the *Larousse Encyclopedia of Astronomy*, two books that I grew up with, kindling my own interest in astronomy. Brilliantly illustrated, every page of *Skywatching* is meticulously laid out and blended into a single volume that will serve as a reference manual in the library as well as a field guide at the telescope. All involved in the project deserve credit for creating the best popular book on amateur astronomy that I have seen in a long time. □

Barry Madore, the father of a seven- and an eight-year old, is at IPAC, California Institute of Technology, 770 South Wilson, Pasadena, California 91125, USA.

LITTLE ONES

More light than heat

Ros Cotter

Fire: Friend or Enemy? Kingfisher, £12.99; published as *Taming Fire* in the USA by Scholastic at \$19.95. The cover entices the young reader in through an impressive embossed display of erupting magma. Fire continues to play visually upon the topics inside, which besides volcanoes include lightning, cooking, blacksmithing, rockets and fireworks. There are plenty of historical snippets too (on firearms, fire festivals and signalling with fire, for example) and a whole cast of celebrities (Thor, Prometheus, Lavoisier, Nobel and Franklin, to name a few).

The language is lively (that magnificent word 'conflagration' is vividly explained) and illuminates an imaginative range of scientific and cultural information. I asked my inquisitive ten-year-old nephew Freddie for his opinion. He would have liked more on lightning and something on the ecological effects of bush fires. Rather disappointingly, he considered that the book belonged in the school library ("for fire projects"), indicating that despite stickers and even a pull-out pizza in the cooking section, it had not quite made it over the education-entertainment barrier.

Taming Fire is indeed comprehensively indexed and is a valuable reference book. But details that might intrigue a parent or teacher (how the tinder mushroom works, why humans started to cook their food) kindled less interest in Freddie than the fact that tallow candles "filled the room with smoke and smelled awful".

The book ends with a rather drab page of fogeyish safety tips (the reader is advised to encourage his or her parents to check smoke detectors regularly). Given the dangerous fascination and excitement of fire, perhaps these would have benefited from the interactive treatment accorded to other subjects. □

Ros Cotter is on the editorial staff of *Nature*.

Energetic reading

Ian Fells

From the Big Bang to Electricity. Kingfisher, £12.99; published as *Exploring Energy* in the USA by Scholastic at \$19.95. This is a very colourful book with lively diagrams, photographs and drawings printed on a black glossy background. It deals with energy in an encyclopaedic way, starting with the Sun and proceeding through muscles, water wheels, steam, electricity, light bulbs (first invented by Swan, not Edison as the text claims) all the way to nuclear power and renewables.

A certain amount of 'do it yourself' is

required of the reader; diagrams and drawings open out and parts of pages are removed by tearing along dotted lines. I was reluctant to let my young grandson tear bits of pages out of a book; not a habit to be encouraged. But he enjoyed applying stickers of rockets, engines, nuclear steam presurisers (*sic*) and bicycle dynamos to the appropriate spaces on the pages. We liked the real feather overlaid by a plastic sheet that becomes charged when the plastic is rubbed and then removed.

The detail in the book is considerable, with a strong historical gloss only really appropriate for 11-12-year-olds. The price for 45 pages seems a bit expensive, even for an indulgent grandfather. □

Ian Fells is in the Department of Chemical Engineering, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.

Change for the better

Trevor Davies

The World of Weather. Kingfisher, £12.99; published as *Wind and Weather* in the USA



by Scholastic at \$19.95. I do not often read children's books. Perhaps I should, as I enjoyed this one: the snappy design meant that I could dip in and extract fascinating information before my attention span was exhausted and allowed me to indulge my indiscipline of reading pages in an almost random order.

There are things to do, such as opening up inner pages, exploring the overlays, using 3D glasses (impressive!) and placing stickers on pages. The pages are colourful and 'busy', but the science is presented in eminently digestible pieces.

The weather is, of course, a wonderfully accessible medium for illustrating physics and the book takes full advantage of this. I would like to think that the young reader will absorb some of the excitement of science that the book conveys. The description and explanations for weather, and the underlying physics, are sound, although the pedant would argue that a few have been compromised in the cause of explanation. The publishers are justified in their approach: the important thing is to interest a child sufficiently for him or her to want to know more. Only the 'targeted consumer' can be a judge in that regard, but this middle-aged 'expert' believes that they have put together an entertaining and informative "living encyclopedia". I will be disappointed if it does not succeed. □

Trevor Davies is in the Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK.